

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 83 Seconds

(without alignments)  
260.082 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678

Sequence: 1 MARKSGDIEKIKKLVLI.....LAKPLTKDKIILINQLMDA 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	100.0	136	20	AA1980
2	678	100.0	136	20	AA1981
3	538	79.4	142	21	AA1982
4	538	79.4	142	21	AA1983
5	515	76.0	142	20	AA1984
6	472	69.6	116	20	AA1985
7	437	64.5	104	21	AA1986
8	437	64.5	104	21	AA1987
9	301	44.4	70	21	AA1988

10	280	41.3	56	21	AA1989
11	213	31.4	71	21	AA1990
12	202	29.8	251	21	AA1991
13	192.5	28.4	67	20	AA1992
14	171.5	25.3	118	23	AA1993
15	156.5	23.1	471	24	AA1994
16	156.5	23.1	496	24	AA1995
17	156.5	23.1	666	24	AA1996
18	156.5	23.1	709	24	AA1997
19	156.5	23.1	719	24	AA1998
20	156.5	23.1	802	24	AA1999
21	156.5	23.1	870	24	AA2000
22	156.5	23.1	874	24	AA2001
23	156.5	23.1	887	24	AA2002
24	156.5	23.1	900	24	AA2003
25	156.5	23.1	906	24	AA2004
26	156.5	23.1	931	24	AA2005
27	155	22.9	922	22	AA2006
28	149	22.0	1373	23	AA2007
29	148.5	21.9	922	22	AA2008
30	148.5	21.9	922	22	AA2009
31	148.5	21.9	950	23	AA2010
32	146	21.5	1018	21	AA2011
33	141	20.8	1146	20	AA2012
34	139.5	20.6	1281	19	AA2013
35	139.5	20.6	1298	19	AA2014
36	139.5	20.6	1298	20	AA2015
37	139.5	20.6	1298	20	AA2016
38	137	20.2	125	23	AA2017
39	134.5	19.8	1122	23	AA2018
40	134	19.8	120	22	AA2019
41	133.5	19.7	1023	22	AA2020
42	131	19.3	136	23	AA2021
43	131	19.3	674	21	AA2022
44	130	19.2	1081	20	AA2023
45	130	19.2	1081	20	AA2024

ALIGNMENTS

RESULT 1

AA1980  
ID AAY42637 standard; Protein; 136 AA.

AC AAY42637;

DT 10-JAN-2000 (first entry)

DE Brassica response regulator protein D22.

KW Signal transduction protein; dehiscence; male sterile plant; D22 gene; shatter resistance; oilseed rape; response regulator protein.

XX Brassica napus.

XX WO9949046-A1.

PD 30-SEP-1999.

XX 22-MAR-1999; 99WO-CB00905.

XX 20-MAR-1998; 98GB-0006113.

XX (BIOG-) BIOEMMA UK LTD.

XX Wyatt P, Roberts JA, Whitelaw C;

XX WPI; 1999-580449/49.

XX N-PSDB; AA222974.

XX A nucleic acid encoding a signal transduction protein involved in plant dehiscence, useful for producing shatter resistant male sterile plants



PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139432.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 29-JUN-1999; 99US-0141287.  
PR 30-JUN-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0141844.  
PR 01-JUL-1999; 99US-0141854.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0148171.  
PR 10-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148341.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150568.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152263.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161559.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.

```
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.4%; Score 538; DB 21; Length 142;
Best Local Similarity 78.9%; Pred. No. 4.7e-52;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 1 MATKSMGDIK-----IKKKL-NVLIVDDPLNLIHEKIIKAIGGISOTANNGEEAVII 54
    |||||  ||  :||| ||||| ||||| :|| ||||| ||||| ||||| :
DB 1 MATKSTGGTEKTSIEVKKKLNLVIVDDPLNRLHEMIKTIKGGISQAKNGEEAVIL 60
    |||||  ||  :||| ||||| ||||| :|| ||||| ||||| ||||| :

QY 55 HRDGGSFLLMDKEMPERDGVSTTKKLREMEVKSIMVGVTSADNEERRAFMEAGLN 114
    ||||| :||| ||||| ||||| ||||| : ||||| ||||| ||||| :
DB 61 HRDGEASFDLLMDKEMPERDGVSTTKKLREMKVTSIMVGVTSVADQEEERKAFMEAGLN 120
    ||| ||||| ||||| ||||| : ||||| ||||| ||||| :

QY 115 HCLAKPLTKDKIPLINQLMDA 136
    ||| ||||| ||||| ||||| : ||||| ||||| ||||| :
DB 121 HCLEKPLTKAKIFPLISHLFDA 142
    ||| ||||| ||||| ||||| : ||||| ||||| ||||| :

RESULT 4
AAG44439
ID AAG44439 standard; Protein; 142 AA.
XX
AC AAG44439;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55668.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132046.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144894.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145056.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.4%; Score 538; DB 21; Length 142;
Best Local Similarity 78.9%; Pred. No. 4.7e-52;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 1 MATKSMGDIK-----IKKL-NVLIYDDPLNLITHEKIKAIGGISQTANNGEAVII 54
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MATKSTGGTEKTSIEVKKKLINVLIVDDDP LNRRLEHMIKIITIGGISQTAKNGEAVIL 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 55 HRDGGSEFDLILMDKEMPERGVSTTKKLEMEVKSMIVGVTSIADNEEERRAFMEAGLN 114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 HRDGEASFDLILMDKEMPERGVSTTKKLEMKVTSIVGVTSVADQEEERKAFMEAGLN 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 115 HCLAKPLTKDKIIPLIINOLMDA 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 HCLERPLTKAKIFPLISHLEDA 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
AAY42645
ID AAY42645 standard; Protein; 142 AA.
XX AC AAY42645;
XX DT 10-TAN-2000 (first entry)
XX DE A. thaliana DZ2AT3 putative peptide sequence.
XX KW Signal transduction protein; dehiscence; male sterile plant;
    shatter resistance; oilseed rape; DZ2AT3 gene.
XX OS Arabidopsis thaliana.
XX FH Key Location/Qualifiers
FT Misc-difference 57 /note= "unknown"
FT FT /note= "unknown"
FT FT /note= "unknown"
XX PN WO9949046-A1.
XX PD 30-SEP-1999.
XX PF 22-MAR-1999; 99WO-GB00905.
XX PR 20-MAR-1998; 98GB-0006113.
XX PA (BIOG-) BIOGEMMA UK LTD.
XX PI Wyatt P, Roberts JA, Whitelaw C;
XX WPI; 1999-580449/49.
XX N-PSDB; AAZ22978.
XX A nucleic acid encoding a signal transduction protein involved in plant
    dehiscence, useful for producing shatter resistant male sterile plants
    -
XX Example 3; Fig 9; 7lpp; English.
XX The invention provides a nucleic acid encoding a signal transduction
    protein involved in the process of dehiscence. The nucleic acids and
    proteins are useful for regulating or controlling dehiscence of a pod or
    an anther in a plant, useful in the production of male sterile plants.
    The methods, etc. may be used in production of shatter resistance or
    shatter-delayed plants such as oilseed rape (Brassica napus). The present
    sequence represents a DZ2 A. thaliana homologue DZ2AT3 putative peptide
    sequence.
    CC 99US-0161359.
```



100

Db 2 IIKTIIGISQAKNGEAVILHRDGEASFDLILMDKEMPERDGVSTTKKLRMKVTSIV 61  
QY 94 GVTSLADNEERAEAFWEAGLNHCLAKPLTKDKIIPILINOLMDA 136  
Db 62 GVTSDAQDEERAEAFWEAGLNHCLAKPLTKDKIIPILINOLMDA 104

RESULT 8  
ID AAG44440 standard; Protein; 104 AA.  
AC AAG44440;  
XX  
DT 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55669.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0123788.  
PR 25-MAR-1999; 99US-0126264.  
PR 28-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127452.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 28-APR-1999; 99US-0130510.  
PR 30-APR-1999; 99US-0130591.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132563.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146366.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154339.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 18-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159684.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
 PR 29-OCT-1999; 99US-0162143.

Query Match 54.5%; Score 437; DB 21; Length 104;  
 Best Local Similarity 84.5%; Pred. No. 5.9e-41;  
 Matches 87; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 34 I I K A I G I S T A N G E A V I I H R D G S S F D L I I M D K E M P E R D G V S T T K L R E M E V K S N I V 93  
 DQ 2 I I K T I G I S T A N G E A V I I H R D G S S F D L I I M D K E M P E R D G V S T T K L R E M E V K S N I V 61  
 QY 94 G V T S L A D N E E E R A F R M E A G L N H C L K P L T K D K I I P L I N O L M D A 136

Db 62 G V T S V A D Q E E R R A F M E A G L N H C L K P L T K A K I F P L I S H L F D A 104  
 RESULT 9  
 AAG21089  
 ID AAG21089 standard; Protein: 70 AA.  
 XX  
 AC AAG21089;  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 23518.  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.



RESULT 10  
AAG21090  
ID AAG21090 standard; Protein: 66 AA.  
XX AC AAG21090;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 23519.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134766.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0137822.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 08-AUG-1999; 99US-0147303.  
PR 08-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR	20-AUG-1999;	99US-0149723.	ID	XX	AG21088 standard; Protein; 71 AA.
PR	20-AUG-1999;	99US-0149929.	XX	AC	AG21088;
PR	23-AUG-1999;	99US-0149902.	XX	XX	17-OCT-2000 (first entry)
PR	23-AUG-1999;	99US-0149930.	DT	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 23517.
PR	25-AUG-1999;	99US-0150566.	DE	XX	Protein identification; signal transduction pathway; metabolic pathway;
PR	26-AUG-1999;	99US-0150884.	DE	XX	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151065.	DE	XX	termination sequence.
PR	27-AUG-1999;	99US-0151066.	DE	XX	Arabidopsis thaliana.
PR	27-AUG-1999;	99US-0151080.	DE	XX	EP1033405-A2.
PR	30-AUG-1999;	99US-0151303.	DE	XX	06-SEP-2000.
PR	31-AUG-1999;	99US-0151438.	DE	XX	25-FEB-2000; 2000EP-0301439.
PR	01-SEP-1999;	99US-0151930.	DE	XX	25-FEB-1999; 99US-0121825.
PR	01-SEP-1999;	99US-0152363.	DE	XX	05-MAR-1999; 99US-0123180.
PR	10-SEP-1999;	99US-0153070.	DE	XX	09-MAR-1999; 99US-0123548.
PR	13-SEP-1999;	99US-0153758.	DE	XX	23-MAR-1999; 99US-0125788.
PR	16-SEP-1999;	99US-0154018.	DE	XX	25-MAR-1999; 99US-0126264.
PR	16-SEP-1999;	99US-0154039.	DE	XX	01-APR-1999; 99US-0126785.
PR	20-SEP-1999;	99US-0154779.	DE	XX	01-APR-1999; 99US-0127462.
PR	22-SEP-1999;	99US-0155139.	DE	XX	06-APR-1999; 99US-0128234.
PR	23-SEP-1999;	99US-0155486.	DE	XX	08-APR-1999; 99US-0128714.
PR	24-SEP-1999;	99US-0155659.	DE	XX	16-APR-1999; 99US-0129845.
PR	28-SEP-1999;	99US-0156458.	DE	XX	19-APR-1999; 99US-0130077.
PR	29-SEP-1999;	99US-0156596.	DE	XX	21-APR-1999; 99US-0130449.
PR	04-OCT-1999;	99US-0157117.	DE	XX	23-APR-1999; 99US-0130510.
PR	05-OCT-1999;	99US-0157753.	DE	XX	23-APR-1999; 99US-0130891.
PR	06-OCT-1999;	99US-0157865.	DE	XX	28-APR-1999; 99US-0131449.
PR	07-OCT-1999;	99US-0158029.	DE	XX	30-APR-1999; 99US-0132048.
PR	08-OCT-1999;	99US-0158232.	DE	XX	30-APR-1999; 99US-0132407.
PR	12-OCT-1999;	99US-0158369.	DE	XX	04-MAY-1999; 99US-0132484.
PR	13-OCT-1999;	99US-0159293.	DE	XX	05-MAY-1999; 99US-0132485.
PR	13-OCT-1999;	99US-0159294.	DE	XX	06-MAY-1999; 99US-0132486.
PR	13-OCT-1999;	99US-0159295.	DE	XX	06-MAY-1999; 99US-0132487.
PR	14-OCT-1999;	99US-0159329.	DE	XX	07-MAY-1999; 99US-0132863.
PR	14-OCT-1999;	99US-0159330.	DE	XX	11-MAY-1999; 99US-0134256.
PR	14-OCT-1999;	99US-0159331.	DE	XX	14-MAY-1999; 99US-0134218.
PR	14-OCT-1999;	99US-0159637.	DE	XX	14-MAY-1999; 99US-0134219.
PR	18-OCT-1999;	99US-0159638.	DE	XX	14-MAY-1999; 99US-0134221.
PR	21-OCT-1999;	99US-0160741.	DE	XX	14-MAY-1999; 99US-0134370.
PR	21-OCT-1999;	99US-0160767.	DE	XX	18-MAY-1999; 99US-0134768.
PR	21-OCT-1999;	99US-0160768.	DE	XX	19-MAY-1999; 99US-0134941.
PR	21-OCT-1999;	99US-0160770.	DE	XX	20-MAY-1999; 99US-0135124.
PR	21-OCT-1999;	99US-0160814.	DE	XX	21-MAY-1999; 99US-0135353.
PR	21-OCT-1999;	99US-0160815.	DE	XX	21-MAY-1999; 99US-0135629.
PR	22-OCT-1999;	99US-0160980.	DE	XX	25-MAY-1999; 99US-0136021.
PR	22-OCT-1999;	99US-0160981.	DE	XX	27-MAY-1999; 99US-0136392.
PR	23-OCT-1999;	99US-0160989.	DE	XX	28-MAY-1999; 99US-0136782.
PR	25-OCT-1999;	99US-0161404.	DE	XX	01-JUN-1999; 99US-0137222.
PR	25-OCT-1999;	99US-0161405.	DE	XX	03-JUN-1999; 99US-0137528.
PR	25-OCT-1999;	99US-0161406.	DE	XX	04-JUN-1999; 99US-0137502.
PR	26-OCT-1999;	99US-0161359.	DE	XX	07-JUN-1999; 99US-0137724.
PR	26-OCT-1999;	99US-0161360.	DE	XX	08-JUN-1999; 99US-0138094.
PR	26-OCT-1999;	99US-0161361.	DE	XX	10-JUN-1999; 99US-0138540.
PR	28-OCT-1999;	99US-0161920.	DE	XX	10-JUN-1999; 99US-0138647.
PR	28-OCT-1999;	99US-0161922.	DE	XX	14-JUN-1999; 99US-0139119.
PR	28-OCT-1999;	99US-0161993.	DE	XX	16-JUN-1999; 99US-0139452.
PR	29-OCT-1999;	99US-0162142.	DE	XX	16-JUN-1999; 99US-0139453.
PR	29-OCT-1999;	99US-0162142.	DE	XX	17-JUN-1999; 99US-0139454.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139455.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139456.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139457.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139458.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139459.
PR	29-OCT-1999;	99US-0162142.	DE	XX	18-JUN-1999; 99US-0139460.

Query Match 41.3%; Score 280; DB 21; Length 66;  
 Best Local Similarity 84.8%; Pred. No. 1.le-23;  
 Matches 56; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 71 MPEDGVSTTKLRMEVKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIIPLI 130  
 |||||  
 Db 1 MPEDGVSTTKLRMEVKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIIPLI 60  
 |||||

QY 131 NQLMDA 136  
 : |||  
 Db 61 SHLFDA 66

PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0143977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150586.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155138.  
PR 23-SEP-1999; 99US-0155480.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 31.4%; Score 213; DB 21; Length 71;  
Best Local Similarity 73.4%; Pred. No. 3.7e-16;  
Matches 47; Conservative 3; Mismatches 8; Indels 6; Gaps 2;

Qy 1 MATKSMGDIETK-----IKKKL-NVLIVDDPLNLIHEKIKAIIGISQATANGGEAVII 54  
Db 1 MATKSTGGTEKYSIEVKKKLINVLIVDDPLNRRHEMIKTIIGISQATANGGEAVIL 60  
Qy 55 HRDG 58  
Db 61 HRDG 64

RESULT 12  
AAB25159  
ID AAB25159 standard; Protein: 261 AA.  
XX  
AC AAB25159;

XX 27-NOV-2000 (first entry)  
DE Pinus radiata cell signalling involved protein SEQ ID NO:127.  
DE  
DE  
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KW environmental change; development; cell proliferation; differentiation;  
KW elongation; survival; disease resistance; nutrient metabolism.  
XX  
OS Pinus radiata.  
XX  
PN WO200042171-A1.  
XX  
XX 20-JUL-2000.  
XX  
XX 11-JAN-2000; 2000WO-US00724.  
XX  
XX 12-JAN-1999; 99US-0228986.  
XX  
XX 01-NOV-1999; 99US-0162866.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Strabala TJ, Nieuwenhuizen NJ;  
XX  
XX WPI: 2000-476052/41.  
XX  
XX Isolated polynucleotide encoding a polypeptide involved in cell  
XX signaling used for generating transgenic plants with modified responses  
XX to external signals -  
XX  
XX Claim 3; Page 104-105; 527pp; English.  
XX  
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
XX and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
XX pine (Pinus radiata) also known as Monterey pine). The protein sequences  
XX are involved in cell signalling. The polynucleotide and protein  
XX sequences can be used to modify the response of plant cells to external  
XX signals e.g. environmental changes or pathogens during the growth and  
XX development of a plant. They can be used to modify cell proliferation,  
XX differentiation, elongation and survival, resistance to disease and  
XX nutrient metabolism. Examples of modifications which can be produced are  
XX altered fruit ripening and senescence of leaves and flowers e.g. to  
XX delay senescence and prolong the life of cut flowers or enhance  
XX senescence of reproductive organs to engineer sterile plants. Other  
XX modifications can be used to delay senescence in selected cell types or  
XX organs providing fruit and vegetables which have a longer shelf life  
XX between harvest and consumption, or to decrease branching frequency in  
XX forest tree species giving long stretches of valuable knot-free clear  
XX wood which can be used in solid timber furniture and veneers.  
XX  
XX Sequence 261 AA;  
XX  
Query Match 29.8%; Score 202; DB 21; Length 261;  
Best Local Similarity 36.5%; Pred. No. 3.7e-14;  
Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;  
QY 18 VLIVDDPLNLIIEHKIIKAGIGISOTANGBEAVIIHRDGSFSLILMDKEMPERDGV 77  
Db 137 ILIVDDTQINRIIFGRVLOSILNLYCEAENGKAVDYFKQ -GRTYDLVLMKEMPVMDGH 195  
QY 78 STTKKLEMEVKSMIVGVTSLADNEERAFWEAGLNHCLAKPLKDKIILPINO 132  
Db 196 EATFQLRSMGVKTPIVALTAA-NTLQSDKDLFFEAGVDYDFQSKPLSRDLVLQDQ 249  
RESULT 13  
ID AAY42644  
XX AAY42644 standard; Protein; 67 AA.  
XX  
XX AAY42644;  
XX  
XX 10-JAN-2000 (first entry)  
DT

XX Brassica napus DZ2B partial fragment.  
DE  
XX Signal transduction protein; dehiscence; male sterile plant; DZ2B gene;  
KW shatter resistance; oilseed rape; response regulator protein.  
XX  
XX Brassica napus.  
XX  
XX WO9949046-A1.  
XX  
XX 30-SEP-1999.  
XX  
XX 22-MAR-1999; 99WO-GB009005.  
XX  
XX 20-MAR-1998; 98GB-0006113.  
XX  
XX (BIOG-) BIOGEMMA UK LTD.  
XX  
XX Wyatt P, Roberts JA, Whitelaw C;  
XX  
XX WPI: 1999-580449/49.  
XX  
XX N-PSDB; AAZ22975.  
XX  
XX A nucleic acid encoding a signal transduction protein involved in plant  
XX dehiscence, useful for producing shatter resistant male sterile plants  
XX -  
XX  
XX Example 2; Fig 5; 71pp; English.  
XX  
XX The invention provides a nucleic acid encoding a signal transduction  
XX protein involved in the process of dehiscence. The nucleic acids and  
XX proteins are useful for regulating or controlling dehiscence of a pod or  
XX an anther in a plant, useful in the production of male sterile plants.  
XX The methods, etc. may be used in production of shatter resistance or  
XX shatter-delayed plants such as oilseed rape (Brassica napus). The present  
XX sequence represents the partial fragment of B. napus DZ2B protein.  
XX  
XX Sequence 67 AA;  
XX  
Query Match 28.4%; Score 192.5; DB 20; Length 67;  
Best Local Similarity 65.7%; Pred. No. 6.7e-14;  
Matches 44; Conservative 6; Mismatches 10; Indels 7; Gaps 2;  
QY 1 MAT--KSMGDIEK-----IKKKLNVLIVDDPLNLIIEHKIIKAGIGISOTANGBEAVI 53  
Db 1 MATTTSTGDIKKTKRSVEVKKLNVLIVDDDTVIRKLHENIKISIGISOTAKNGEEAVN 60  
QY 54 IHRDGS 60  
Db 61 IHRDGN 67  
RESULT 14  
ABG70785  
ID ABG70785 standard; Protein; 118 AA.  
XX  
XX ABG70785;  
XX  
XX 09-DEC-2002 (first entry)  
XX  
XX DE E. coli RscC receiver region of histidine kinase.  
XX  
XX RscC; antagonist; agonist; cytokinin receptor;  
KW receptor; signal transduction; histidine kinase; hormone; cell division;  
KW cell differentiation; agriculture; growth regulator; harvest.  
XX  
XX Escherichia coli.  
XX  
XX EPI241182-A2.  
XX  
XX 18-SEP-2002.  
XX  
XX 13-MAR-2002; 2002EP-0005749.  
XX  
XX

XX 15-MAR-2001; 2001JP-0073812.  
PR 29-JUN-2001; 2001JP-0198639.  
PR 29-JUN-2001; 2001JP-0198640.  
XX (SUMO ) SUMITOMO CHEM CO LTD.  
XX Kakimoto T, Higuchi M, Inoue T;  
XX WPI: 2002-693041/75.  
XX  
XX Analyzing agonist or antagonist activity of a substance for use as a  
PT plant growth regulator, comprises measuring intracellular signal  
PT transduction from a cytokinin receptor expressed in a cell contacted  
PT with the test substance.  
XX  
XX Disclosure; Page 42; 47pp; English.  
XX  
XX The invention discloses a method for analysing antagonist or agonist  
CC activity to a cytokinin receptor. The method comprises bringing a  
CC candidate substance into contact with a transformed cell, in which a DNA  
CC encoding the receptor has been introduced, and then measuring the  
CC existence, or the quantity, of the intracellular signal transduction from  
CC the receptor expressed in the cell. The cytokinin receptor comprises an  
CC extracellular region of the receptor, transmembrane regions, a histidine  
CC kinase region and a receiver region of the kinase. The transmembrane  
CC regions and kinase region are homogeneous to each other and the receptor  
CC region is heterogeneous to them. Cytokinins are plant hormones relevant  
CC to cell division and differentiation of higher plants. The method is used  
CC for analysing agonist or antagonist activity to a cytokine receptor. A  
CC substance with agonist or antagonist activity to the receptor can be  
CC used, in agriculture, as a plant growth regulator, e.g. after harvest.  
CC The advantage is that the candidate substances do not need to be prepared  
CC in such large amounts as in previous methods and that the method avoids  
CC the immensely long time to observe and evaluate the growth of the plant  
CC and the physiological changes of the plant after spraying. The sequence  
CC presented is the E. coli receiver region of histidine kinase which can  
CC transmit signals to the cytokinin receptor.  
XX  
XX Sequence 118 AA;  
SQ  
Query Match 25.3%; Score 171.5; DB 23; Length 118;  
Best Local Similarity 33.9%; Pred. No. 3.2e-11;  
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;  
QY 18 VLIVDDPLNLIIEHKIIKAIGISQTANGGEAVIIHRDGGSSFDLILMDKEMPERDG 77  
Db 6 ILVVDHPINRLLADQLGSLGYQCKTANGDYDALNLV--SKNHDIIVLSDVNNPMDGY 63  
QY 78 STTKKLEMEVKSMIVGVTSLADNEERAFMEAGLNHCLAKPLTKDKI 126  
Db 64 RLTORIRQLGLTLFVIGVTANALAEKQRC-LESGMDSCLSKRPVTLDWI 111  
RESULT 15  
AAO26983  
ID ARO26983 standard; Protein; 471 AA.  
XX AAO26983;  
XX  
XX 10-MAY-2003 (first entry)  
XX Pseudomonas aeruginosa PvrR related protein, SEQ ID NO 4.  
XX Antibacterial; PvrR; variant Pseudomonas; microorganism; gram negative;  
KW phenotype-mediated antibiotic-resistance; gram-positive;  
KW bacterial infection.  
XX Pseudomonas aeruginosa.  
OS  
XX WO2003004691-A2.  
XX  
XX 16-JAN-2003.  
PD

XX 05-JUL-2002; 2002WO-US23242.  
XX  
XX 06-JUL-2001; 2001US-303286P.  
PR 16-APR-2002; 2002US-37233P.  
XX  
XX (GEO ) GEN HOSPITAL CORP.  
XX Ausubel FM, Drenkard E;  
XX  
XX WPI: 2003-221608/21.  
DR N-PSDB; AAL53304.  
XX  
XX New isolated PvrR polypeptide and polynucleotide that regulates  
PT bacterial biofilm formation, useful for the diagnosis, prevention and  
PT treatment of gram-negative or gram-positive bacterial infection.  
XX  
XX Disclosure; Fig 5E; 185pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising a PvrR  
CC (variant Pseudomonas) amino acid sequence having at least 50 % identity  
CC to a 399 residue amino acid sequence, given in the specification, where  
CC expression of the polypeptide, in a microorganism, affects phenotype-  
CC mediated antibiotic-resistance in the microorganism. The methods and  
CC compositions of the present invention are useful for the diagnosis,  
CC prevention and treatment of gram negative or gram-positive bacterial  
CC infection. This sequence represents a Pseudomonas protein used in the  
CC exemplification of the invention.  
XX  
XX Sequence 471 AA;  
SQ  
Query Match 23.1%; Score 156.5; DB 24; Length 471;  
Best Local Similarity 35.8%; Pred. No. 9.9e-09;  
Matches 38; Conservative 22; Mismatches 43; Indels 3; Gaps 3;  
QY 16 LNVIVDDPLNLIIEHKIIKAIGISQTANGGEAVIIHRDGGSSFDLILMDKEMPERD 75  
Db 352 LRVVVVEDNAILQLIRQMEALGCVSVELLDGREA-LLHCQ-TACFDVVLTDINMPNAN 409  
QY 76 GVSTTKKLEMEVKSMIVGVTSLADNEERAFMEAGLNHCLAKPL 121  
Db 410 GYELTAELRRQGFQPIIGATANAMREERERC-MSAGMNDCLVKPV 454  
Search completed: August 14, 2003, 16:50:31  
Job time : 84 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 29 Seconds  
(without alignments)  
198.423 Million cell updates/sec

Title: US-09-646-679-15  
Perfect score: 678  
Sequence: 1 MATKSMGDIKIKKLNVL.....LAKPLTKDKIPLINQLMDA 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202	29.8	261	US-09-228-986-127	Sequence 127, App
2	152.5	22.5	946	US-09-328-352-7973	Sequence 7973, Ap
3	150	22.1	1014	US-09-252-991A-31338	Sequence 31338, A
4	148	21.8	947	US-09-252-991A-29359	Sequence 29359, A
5	147	21.7	1627	US-09-252-991A-20395	Sequence 20395, A
6	145	21.4	1220	US-08-843-530B-36	Sequence 36, Appl
7	144	21.2	162	US-09-252-991A-17557	Sequence 17557, A
8	139.5	20.6	1281	US-08-843-530B-6	Sequence 6, Appl
9	139.5	20.6	1298	US-08-843-530B-2	Sequence 2, Appl
10	139.5	20.6	1298	US-08-843-530B-4	Sequence 4, Appl
11	139.5	20.6	1298	US-08-843-530B-34	Sequence 34, Appl
12	139.5	20.6	1441	US-09-252-991A-28143	Sequence 28143, A
13	138.5	20.4	234	US-09-634-238-220	Sequence 220, App
14	135.5	20.0	760	US-09-252-991A-25928	Sequence 25928, A
15	135.5	20.0	860	US-09-252-991A-26112	Sequence 26112, A
16	133.5	19.7	918	US-08-843-530B-35	Sequence 35, Appl
17	131	19.3	764	US-09-252-991A-18607	Sequence 18607, A
18	131	19.3	1036	US-09-252-991A-27075	Sequence 27075, A
19	130	19.2	1081	US-08-843-530B-18	Sequence 18, Appl
20	130	19.2	1117	US-08-843-530B-33	Sequence 33, Appl
21	127	18.7	506	US-09-252-991A-18168	Sequence 18168, A
22	125	18.4	227	US-09-107-532A-6769	Sequence 6769, Ap
23	124.5	18.4	133	US-09-328-352-5134	Sequence 5134, Ap
24	124	18.3	971	US-09-112-450-2	Sequence 2, Appl
25	124	18.3	971	US-09-419-291A-2	Sequence 2, Appl
26	124	18.3	2471	US-09-112-450-4	Sequence 4, Appl
27	124	18.3	2471	US-09-419-291A-4	Sequence 4, Appl

28	122.5	18.1	265	4	US-09-328-352-6073	Sequence 6073, Ap
29	122	18.0	311	4	US-09-328-352-991A-22932	Sequence 22932, A
30	120.5	17.8	811	4	US-09-252-991A-22216	Sequence 22216, A
31	119	17.6	484	4	US-09-252-991A-31677	Sequence 31677, A
32	117	17.3	320	4	US-09-328-352-6809	Sequence 6809, Ap
33	115.5	17.0	331	4	US-09-252-991A-23765	Sequence 23765, A
34	115	17.0	762	4	US-09-228-986-114	Sequence 114, App
35	113	16.7	302	4	US-09-252-991A-29564	Sequence 29564, A
36	112.5	16.6	129	4	US-09-252-991A-16981	Sequence 16981, A
37	112.5	16.6	659	4	US-09-252-991A-17904	Sequence 17904, A
38	111.5	16.4	212	4	US-09-107-532A-6859	Sequence 6859, Ap
39	111.5	16.4	234	4	US-09-107-532A-6746	Sequence 6746, Ap
40	110.5	16.3	256	4	US-09-328-352-5430	Sequence 5430, Ap
41	110	16.2	599	4	US-09-228-986-117	Sequence 117, App
42	109	16.1	245	4	US-09-134-001C-3779	Sequence 3779, Ap
43	109	16.1	257	4	US-09-328-352-7397	Sequence 7397, Ap
44	108.5	16.0	246	4	US-09-107-532A-5040	Sequence 5040, Ap
45	108.5	16.0	250	4	US-09-634-238-221	Sequence 221, App

## ALIGNMENTS

### RESULT 1

US-09-228-986-127  
; Sequence 127, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228.986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 127  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-127

Query Match 29.8%; Score 202; DB 4; Length 261;

Best Local Similarity 36.5%; Pred. No. 7e-16;

Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

QY 18 VLIVDDPLNLIHKKIKAIGGISQTANNGEAAVTHRDGGSFLLIMDKEMPERDGV 77

Db 137 ILLVDTQINRIIFGRVLQSLNLYCEAEKGVAVDFKQ-GRTYDLVMDKEMPVMDGH 195

QY 78 STTKLREMEVKSMIVGVTSLADNEERRAEAGLNHCLAKPLTKKIIPLNQ 132

Db 196 EATQRLSRMGVATPIVALTA-NTLQSDKDLFEAGVDDFQSKPLSRDLVQLLQD 249

### RESULT 2

US-09-328-352-7973

; Sequence 7973, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Berton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328.352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7973

; LENGTH: 946

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7973

```

/ SEQ ID NO 29359
/ LENGTH: 947
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29359

Query Match          21.8%; Score 148; DB 4; Length 947;
Best Local Similarity 28.8%; Pred. No. 1.1e-08;
Matches 34; Conservative 33; Mismatches 43; Indels 8; Gaps 4;

QY      18  VLIIVDDPLNLIHKEIKIKAIGISQTANGGEAV-IHRRDGGSSFDLLIMDKEMPERDG 76
DB      18  LLLVDDPNAPLLVQLLLSDGQYAVDSGYAALEVQRE---REDLVFMDVQPMGMDG 745
QY      77  VSTTKLREVEKSMI--VGVTSLADN--EEERRAFMEAGLNHCLAKPLTKDKLIPLI 130
DB      77  QATEAIRRWEAREVSPVIAITAHALSNEKRALQAGMDVLTLPIDEQQLAQVY 803

RESULT 5
US-09-252-991A-20395
/ Sequence 20395, Application US/09252991A
/ Patent No. 6351795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 20395
/ LENGTH: 1627
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20395

Query Match          21.78%; Score 147; DB 4; Length 1627;
Best Local Similarity 31.9%; Pred. No. 3.3e-08;
Matches 37; Conservative 26; Mismatches 45; Indels 8; Gaps 3;

QY      14  KKLNVLIYDDPLNLIHKEIKIKAIGISQTANGGEAVIHRDGGSSFDLLIMDKEMPE 73
DB      1371  RLQVLVYDDHAVNRQILHQQLSFLGHDEVEAENGLSALNLWH--GQFDMVITDCHMPL 1428
QY      74  RDGVSTYTKKLREMEVKS-----MIVGVTSIADNEERRAFMEAGLNHCLAKPLTKD 124
DB      1429  MSGSDLAIRSIQREERNGEPEPVILGLTADAQPEIERC-IOAGMNECLIRPGLD 1483

RESULT 6
US-08-843-530B-36
/ Sequence 36, Application US/08843530B
/ Patent No. 5939306
/ GENERAL INFORMATION:
/ APPLICANT: Selltrennikoff, Claude
/ APPLICANT: Agnan, Jacqueline
/ APPLICANT: Alex, Lisa A.
/ APPLICANT: Simon, Melvin I.
/ TITLE OF INVENTION: Osmosensing Histidine Kinases
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Medlen & Carroll, LLP
/ STREET: 220 Montgomery Street, Suite 2200
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: United States of America
/ ZIP: 94104
/ COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,530B  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: UTC-02717  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1220 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-843-530B-36

Query Match 21.4%; Score 145; DB 2; Length 1220;  
Best Local Similarity 29.0%; Pred. No. 3.8e-08;  
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;  
QY 8 DIEKIKKLNVLIVDDPLNLIHEKIKA---IGGIS-QTANGGEA---VIHRDGG 59  
DB 1080 DDKNETSVKILVVDHVN---QEVIKMLNLSGIEINELACDQAEFKVLTSG 1135  
QY 60 SFDLILMDKEMPERDGVSTTKL-REMEVKSMIVGVTSLADNEERAFMEAGLNHCLA 118  
DB 1136 ENYMLFMVQMPKVDGLSTMIKRDGLGVTSPVLTAFADDSNIKEC-LESGMNGFLS 1194  
QY 119 KPLTKDKIPLNQLMDA 136  
DB 1195 KPIKRPKLTILFECAA 1212

RESULT 7  
US-09-252-991A-17557  
Sequence 17557, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17557  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17557

Query Match 21.2%; Score 144; DB 4; Length 162;  
Best Local Similarity 30.9%; Pred. No. 2.6e-09;  
Matches 38; Conservative 28; Mismatches 49; Indels 8; Gaps 3;  
QY 18 VLIIVDDPLNLIHEKIKAIGGISQTANGGEAVIHRDGGSSFDLILMDKEMPERDGV 77  
DB 35 ILVAEDNPNVLVGRGLAKRGYAVLAGNRLALDEYLRDPNGIQLILMDGEMPEMDGF 94  
QY 78 STTKKLREME-----VKSMIVGVTSLADNEERAFMEAGLNHCLAKPLTKDKIPLN 131

DB 95 EATRLIRREERAQGWPRVPIVALTALHILD--EHRAGIEAGMDAYLGKPYDRAELVATLE 152  
QY 132 QLM 134  
DB 153 RLL 155  
RESULT 8  
US-08-843-530B-6  
Sequence 6, Application US/08843530B  
Patent No. 5939306  
GENERAL INFORMATION:  
APPLICANT: Selltrennikoff, Claude  
APPLICANT: Aguan, Jacqueline  
APPLICANT: Alex, Lisa A.  
APPLICANT: Simon, Melvin I.  
TITLE OF INVENTION: Osmosensing Histidine Kinases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,530B  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: UTC-02717  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1281 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-843-530B-6

Query Match 20.6%; Score 139.5; DB 2; Length 1281;  
Best Local Similarity 31.1%; Pred. No. 1.8e-07;  
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;  
QY 2 ATKSMGDIKIKKLNVLIVDDPLNLIHEKIIKRAIGGISQTANGGEAV-IHRDGG 60  
DB 1075 ATPSLAD---NTRKSFELLAEADNTVQRLAVKILEKYHHVTVVGNGBEAEVAKR---K 1128  
QY 61 SFDLILMDKEMPERDGVSTTKLREME---VKSMIVGVTSLADNEERAFMEAGLNH 115  
DB 1129 KFDVILMDVQMPINGGFEATAKIREYERSLGSORTPIALTAAH-MMGDREKCIQAOMDE 1187  
QY 116 CLAKPLTKDKIPLI 130  
DB 1188 YLSKPLQNHUQTI 1202

RESULT 9  
US-08-843-530B-2  
Sequence 2, Application US/08843530B  
Patent No. 5939306  
GENERAL INFORMATION:  
APPLICANT: Selltrennikoff, Claude

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
;
US-08-843-530B-4
;
; Query Match 20.6%; Score 139,5; DB 2; Length 1298;
; Best Local Similarity 31.1%; Pred. No. 1.8e-07;
; Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;
;
QY . 2 ATKSMGDIKKKKLNVLIVDDPLNLIIHKIIKAIGGISQTANNGEAV-IHRRDGS 60
; . : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1075 ATPSLAD---NTSFILLAEONTNQLRAVKILEKHYHVTVVYNGGEEAVEAKR---X 1128
; . : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 SFDLITLDMKEMPERDGVSTTKKLREME-----VKSMIVGVTSIADNEETERRAFMEAGLNH 115
; . : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1129 KFDVILMDVQMPGNGFEATKIREYERSUGSRPPIALTAHA-MMGDRKCIQAQWDE 1187
; . : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 CLAKPLTKDKIILPI 130
; . : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1188 YLSKPLQONHLIOTI 1202
; . : : : : : : : : : : : : : : : : : : : : : : : : :
;
RESULT 11
; US-08-843-530B-34
; Sequence 34, Application US/088435530B
; Patent No. 5939306
;
; GENERAL INFORMATION:
; APPLICANT: Selltreinnikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:

```









```
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801.368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 332
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-332

Query Match 21.4%; Score 145; DB 10; Length 1220;
Best Local Similarity 29.0%; Pred. No. 6.6e-07;
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DIEIKKKLNVLIVDDPLNLIHEKIIKAIGGISQTANNGEEA---VTHRDGG 59
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1080 DDDKNETSVKILVDEHNVN---QEVIKRLNLEGIEINELACDGOEAFKVKELTSGK 1135
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 60 SFFDLILMDKEMPERDGVSTTKKL-REMEVKSIMVGVTSIADNEERAFMEAGLNHCLA 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1136 ENYNMIFDVMQPKVDGLLSTKMIRDLGYTSPIVALTAFADDSNIREC-LESNGNGLS 1194
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 119 KPLTKDKIIPINQLMDA 136
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1195 KPIKRPKLTILTERCAA 1212
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-801-368-366
; Sequence 366, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801.368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 366
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae

; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-366

Query Match 21.2%; Score 143.5; DB 10; Length 712;
Best Local Similarity 27.9%; Pred. No. 4.5e-07;
Matches 39; Conservative 28; Mismatches 42; Indels 31; Gaps 3;

QY 10 EKIKKKLNVLIVDDPLNLIHEKIIKAIGGISQTANNGEEAVIIHRDGGSSFDLILMDK 69
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 498 EKVPFKNLVIVEDNVINQAILGSLRKHKISYKLAKNQGEAVNIWKEGG--LHLIFMDL 555
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 70 EMPERDGVSTTKKLREMEVKS-----MIVGVTSIADN 101
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 556 QLPVLSGIEAAKQIDFEKQNGIGIOKLNNSHNSLEKGTSKRFSQAPVIIVALTASNSQ 615
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 102 BEERRAFMEAGLNHCLAKPL 121
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 616 MDKRRKALL-SGNDNDYLTAKPV 634
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-10-126-120-2
; Sequence 2, Application US/10126120
; Publication No. US20030108526A1
; GENERAL INFORMATION:
; APPLICANT: Sakakibara, Hitoshi
; APPLICANT: Takei, Kentaro
; TITLE OF INVENTION: MICROORGANISMS FOR USE IN THE MEASUREMENT OF ENVIRONMENTAL FAC
; FILE REFERENCE: 11127-004001
; CURRENT APPLICATION NUMBER: US/10/126,120
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: JP 2001-291059
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Zea mays
US-10-126-120-2

Query Match 20.3%; Score 137.5; DB 15; Length 974;
Best Local Similarity 24.6%; Pred. No. 3.2e-06;
Matches 35; Conservative 32; Mismatches 46; Indels 29; Gaps 3;

QY 17 NVLIVDDDDPLNLIHEKIIKAIGGISQTANNGEEAVII---HRDGGSSFDLILMDKEMP 72
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 830 NILVVDNKNVLRVAAALKKYGANVSCVESCKDAISLLQPPHR-----FDACFMDVQMP 884
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 73 ERDGVSTTKKLREMEVKSIMVGVTSIADNEEBERRA-----FWEAG 112
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 885 EMDGFATGQIRQEMELKANEERKNKLASIEGSTTAEYHLPVLAMTADVIQATYEKISG 944
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 113 LNHLAKPLTKDKIIPINQLM 134
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 945 MDGYVSKPFDDEQLYQAVSRV 966
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-918-508-7
; Sequence 7, Application US/09918508
; Patent No. US20020177162A1
; GENERAL INFORMATION:
; APPLICANT: KAKIMOTO, TATSUO
; APPLICANT: HIGUCHI, MASAYUKI
; APPLICANT: INOUE, TSUTOMU
; TITLE OF INVENTION: ANALYSIS OF AGONIST-ACTIVITY AND ANTAGONIST-ACTIVITY
; FILE REFERENCE: Q65478
; CURRENT APPLICATION NUMBER: US/09/918,508
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: JP 2001-073812
; PRIOR FILING DATE: 2001-03-15
```

```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-918-508-7

Query Match          20.2%; Score 137; DB 10; Length 125;
Best Local Similarity 29.4%; Pred. No. 2.1e-07;
Matches 37; Conservative 38; Mismatches 37; Indels 14; Gaps 6;

QY 16 LNVIVDDPLNLIIEHKIKA---IGGIS--QTANNGEA---VIIHRDGGSSFDLILM 67
DB 4 VKILVVDNHN---QEVIKMLNLEGIEIELACDGOEAFDKVKELTSKGENYMIFM 59

QY 68 DKMPERDGVSTTKKL-RENEVKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKDKI 126
DB 60 DVQMPKVDGLLSTKMRIRDLGYTSPIVALTAFAADDSNIKEC-LESGMFLSKPRKPKL 118

QY 127 IPLINQ 132
DB 119 KTLITE 124

RESULT 9
US-10-156-761-10049
; Sequence 10049, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10049
; LENGTH: 1829
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10049

Query Match          19.5%; Score 132; DB 15; Length 1829;
Best Local Similarity 31.5%; Pred. No. 3.1e-05;
Matches 39; Conservative 24; Mismatches 51; Indels 10; Gaps 4;

QY 18 LVIVDDPLNLIIEHKIKAIGGISQTANNGEAVII---HRDGGSSFDLILMDKEMPER 74
DB 1710 LVIVDDIRNFALTSVLEQHLVLYAENGREGIEVLEQDHD---VTYVLMIDIMPEM 1765

QY 75 DGVSTTKKLEME--VKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKDKIPLINQ 132
DB 1766 DGVATTATIRMPQFAGLPIIALTAHTAKAMKGDEKA-IESGASDYVKPVPDPDLITVMEQ 1824

QY 133 LMDA 136
DB 1825 WMRA 1828

RESULT 10
US-09-424-951-4
; Sequence 4, Application US/09424951
; Patent No. US20020137034A1
```

```
; GENERAL INFORMATION:
; APPLICANT: SRIKANTHA, THYAGARAJAN
; APPLICANT: SOLL, DAVID R.
; TITLE OF INVENTION: CANDIDA ALBICANS TWO-COMPONENT HYBRID KINASE GENE,
; FILE REFERENCE: 087714/0113
; CURRENT APPLICATION NUMBER: US/09/424,951
; CURRENT FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: PCT/US98/11658
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,914
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-424-951-4

Query Match          18.9%; Score 128; DB 10; Length 1081;
Best Local Similarity 28.3%; Pred. No. 4e-05;
Matches 43; Conservative 31; Mismatches 48; Indels 30; Gaps 7;

QY 5 SMGDIERIKKLNLIYVDDPLNLIIEHKIKAIGGISQTANNGEAV-IIHRDGGSSFD 63
DB 865 SQNSDESRYVK--ILLAEDNLVNQKLAVRLEKQGHVVEVNGLEAYEAIKR---NKYD 919

QY 64 LILMDKEMPERDGVSTTKKLEMEVKS-----MIVGVT'S---LADNEEERRAFME 110
DB 920 VVLMVDQVPMVGSGFEATEKIRQWEKSNPIDSUTFTPIALTAAHMLGDREKS---LA 975

QY 111 AGLNHCLAKPL-----TKDKIPLINQMD 135
DB 976 KGMDYVSKPLKPKLLMQTIKKKCIHNINQLKE 1007

RESULT 11
US-10-135-322-17
; Sequence 17, Application US/10135322
; Publication No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENFEY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AWM
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match          18.8%; Score 127.5; DB 14; Length 2150;
Best Local Similarity 32.5%; Pred. No. 0.00012;
Matches 37; Conservative 20; Mismatches 50; Indels 7; Gaps 3;

QY 13 KKKLNLIYVDDPLNLIIEHKIKAIGGISQTANNGEAVIIHRDGGSSFDLILMDKEMP 72
DB 2024 RKK--ALIYEDNELARKVLAQLFKKIDWTISFAENGREA-LKEITGRCFDIVMDCOMP 2080

QY 73 ERDGVSTTK-----KLREMEVKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLT 122
DB 2081 VLDGFQTTILIRSKERENNNKRMNIVALSGSSSSSFVQDCLDSGMDSGFMGPKIT 2134
```

```
RESULT 12
US-09-815-242-11131
; Sequence 11131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11131
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11131

Query Match 18.7%; Score 126.5; DB 9; Length 227;
Best Local Similarity 29.1%; Pred. No. 6.7e-06;
Matches 34; Conservative 30; Mismatches 48; Indels 5; Gaps 2;

QY 18 VLIYDDDDPLNLIHKEIKKIGGISQANNGEEAVIIHRDGGSSFDLIILMDKEMPERDGV 77
Db 4 LLLYDDDDIETELSTLLELEGEDVETANNGLAAL---QKLNESYKLVLLDVMMPKNGI 60

QY 78 STTKKLRMEVKSIVGVTSIADNEERRAFMEAGLNHCLAKPLTKDKIPLINQML 134
Db 61 ETLKEIR--KVSNPVAMLTARGEDIDRVILGELGADDCFLKPFNDRELIARIKAIL 115

RESULT 13
US-10-156-761-12532
; Sequence 12532, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9975
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12532

Query Match 18.5%; Score 125.5; DB 15; Length 227;
Best Local Similarity 29.2%; Pred. No. 8.6e-06;
Matches 38; Conservative 28; Mismatches 53; Indels 11; Gaps 5;

QY 10 EKIKKLNVLIVDDDDPLNLIHKEIKKIG---GIS--QTANNGEAVIIHRDGGSSFDL 64
Db 4 BESRPARVVVADD---QTVVREGIVMLGLLPGIEVVGAGDGHEAVKLVAE--LNPDV 58

QY 65 ILMDKEMPERDGVSTTKKLRMEVKSIVGVTSIADNEERRAFMEAGLNHCLAKPLTKD 124
Db 59 VLMDLRMPRCDGVEATRRIRAEHPGTQVVVLTYYADESLFPA--LRAGARGYLTKDAGD 117

QY 125 KIPLINQML 134
Db 118 EIVRAVESVL 127

RESULT 15
US-09-815-242-13783
; Sequence 13783, Application US/09815242
```

```
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12532
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12532

Query Match 18.6%; Score 126; DB 15; Length 203;
Best Local Similarity 29.5%; Pred. No. 6.5e-06;
Matches 36; Conservative 20; Mismatches 46; Indels 20; Gaps 3;

QY 17 NVLIYDDDDPLNLIHKEIKKIGGISQ--TANNGEAVIIHRDGGSSFDLIILMDKEMPER 74
Db 3 SVLVCDSDPLAREALRRVATVPGVERVTAANGEE--VIRWAGADRSDLIILMDVMPGL 60

QY 75 DGVSTTKKLRMEVKSIVGVTSIADNEERRAFMEAGLNHCLAKPLTKD 118
Db 61 GGVTVERLLSADPGARIIMLTVAEDLDGVALAAGARGYLHKDASRAELRATVTVQALA 120

QY 119 KP 120
Db 121 DP 122

RESULT 14
US-10-156-761-9975
; Sequence 9975, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9975
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9975

Query Match 18.5%; Score 125.5; DB 15; Length 227;
Best Local Similarity 29.2%; Pred. No. 8.6e-06;
Matches 38; Conservative 28; Mismatches 53; Indels 11; Gaps 5;

QY 10 EKIKKLNVLIVDDDDPLNLIHKEIKKIG---GIS--QTANNGEAVIIHRDGGSSFDL 64
Db 4 BESRPARVVVADD---QTVVREGIVMLGLLPGIEVVGAGDGHEAVKLVAE--LNPDV 58

QY 65 ILMDKEMPERDGVSTTKKLRMEVKSIVGVTSIADNEERRAFMEAGLNHCLAKPLTKD 124
Db 59 VLMDLRMPRCDGVEATRRIRAEHPGTQVVVLTYYADESLFPA--LRAGARGYLTKDAGD 117

QY 125 KIPLINQML 134
Db 118 EIVRAVESVL 127

RESULT 15
US-09-815-242-13783
; Sequence 13783, Application US/09815242
```

```

: Patent No. US20020061569A1
:
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA-011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO 13783
: LENGTH: 232
: TYPE: PRT
: ORGANISM: Salmonella typhi
:
: US-09-815-242-13783

```

Query Match	18.4%	Score 124.5	DB 9	Length 232
Best Local Similarity	28.2%	Pred. 1.1e-05		
Matches	33	Conservative	31	Mismatches 48
				Indels 5
				Gaps 2
QY	18	VLIVDDPLNLI	THEKIKAIGISOTANNGEAVI	IRDGGSSPDILMKEMPERDGV 77
	4	ILLDDRELTSL	KLEMEGNFVAHDDQALELDD	--SIDLLDVMMPKXNGI 60
DB				
QY	78	STTKKLREMEVKS	MIWGVTSLADNEERAFNEAGIN	HCLAKPLTKRIIPLINQLM 134
DB	61	DTLAKRTHOTPYI	--MLTARGNELDRVLGELGADGY	LPKPFNDRELVARIRAIL 135

Search completed: August 14, 2003, 16:59:18  
Job time : 56 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 39 Seconds  
(without alignments)  
335,357 Million cell updates/sec

Title: US-09-646-679-15  
Perfect score: 678  
Sequence: 1 MATKSMGDIKKKLNVL.....LAKPLTKDKIIPLINQLMDA 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	181	26.7	1969	2 T08875	histidine kinase h
2	178.5	26.3	1197	1 G65010	sensor protein evg
3	176.5	26.0	1197	2 A91035	probable sensor fo
4	176.5	26.0	1197	2 C85879	probable sensor fo
5	174	25.7	394	2 H87640	sensor histidine k
6	173.5	25.6	948	2 AD0790	sensor protein Res
7	171.5	25.3	933	1 BVECC	sensor protein res
8	171.5	25.3	933	2 C91017	sensor for ctr cap
9	171.5	25.3	933	2 B85861	hypothetical prote
10	171.5	25.3	957	2 AD0149	two component sens
11	162.5	24.0	642	2 E87644	sensor box histid
12	162.5	24.0	778	2 AD0432	aerobic respiratio
13	159	23.5	117	2 A69487	response regulator
14	159	23.5	736	2 C82151	sensor histidine k
15	159	23.5	1645	2 AG1897	two-component hybr
16	157	23.2	900	2 B7252	sensor box histid
17	156.5	23.1	777	2 A35966	chemotaxis protein
18	155	22.9	690	2 A87580	sensor box histid
19	154.5	22.8	122	2 G69422	response regulator
20	154.5	22.8	572	2 B82198	sensor histidine k
21	153.5	22.6	713	2 A87617	sensor box histid
22	152.5	22.5	120	1 S86645	response regulator
23	152.5	22.5	120	2 G84253	chemotaxis protein
24	152.5	22.5	514	2 A87266	sensor histidine k
25	152	22.4	1014	1 S75023	sensor transducti
26	151.5	22.3	633	2 S75525	sensor transducti
27	150.5	22.2	778	2 AG0906	aerobic respiratio
28	150.5	22.2	939	2 A80032	two-component sens
29	150	22.1	622	1 A49344	cell wall assembly

30 150 22.1 992 2 A83324 probable sensor/re  
31 149.5 22.1 776 1 RGEAR aerobic respiratio  
32 149.5 22.1 778 2 A91140 aerobic respiratio  
33 149.5 22.1 778 2 D85985 aerobic respiratio  
34 148 21.8 925 2 E83529 sensor/response re  
35 147 21.7 1212 2 F83153 probable two-compo  
36 145 21.4 1220 2 S48387 SLN1 protein - yea  
37 144.5 21.3 560 2 E87460 sensor histidine k  
38 144 21.2 857 2 C82424 sensor histidine k  
39 144 21.2 919 2 E83212 probable sensor/re  
40 143.5 21.2 712 2 S64828 SSK1 protein - yea  
41 143 21.1 693 2 C87575 sensor histidine k  
42 142.5 21.0 637 2 D87559 sensory box histid  
43 142.5 21.0 918 2 A10860 sensor protein (im  
44 142 20.9 152 2 H97516 probable response  
45 142 20.9 929 2 B41863 two-component regu

ALIGNMENTS

RESULT 1

T08875 histidine kinase homolog DHKB - slime mold (Dictyostelium discoideum)  
N:Alternate names: Hybrid histidine kinase DHKB  
C:Species: Dictyostelium discoideum  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-May-2000  
C:Accession: T08875  
R:Zinda, M.J.; Singleton, C.K.  
Dev Biol. 196, 171-183, 1998.  
A:Title: The hybrid histidine kinase dhx8 regulates score germination in Dictyostelium  
A:Reference number: Z16506; MUID:98248997; PMID:9576830  
A:Accession: T08875  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1969 <SIN>  
A:Cross-references: EMBL:AF024654; NID:g2460282; PID:g2460283  
A:Experimental source: strain KAX3  
C:Genetics:  
A:Gene: dhx8  
A:Introns: 790/3  
C:Superfamily: response regulator homology  
C:Keywords: Protein kinase; transmembrane protein  
F:1841-1964/Domain: Response regulator homology <RRH>

Query Match 26.7%; Score 181; DB 2; Length 1969;  
Best Local Similarity 35.8%; Pred. No. 5.6e-07;  
Matches 48; Conservative 29; Mismatches 39; Indels 18; Gaps 5;  
QY 10 EKIKKLNVLIVDDPLNLIHEKIIKAIGGISQTANNGEAVIIHRDGGSSFDLILMDK 69  
DB 1836 EKIEK---ILLVDNFVNVKIFSKLLKDSGYIFDVAHNGVEAVECVKKG--AYDLILMDK 1890  
QY 70 EMPERDGVSTTKLREMEVKSMI-----VGVTSLADNP--ERRAFVEAGLNHCL 117  
DB 1891 QMPENDDFEATTAIRELEKSNLIESPFKSHSHVIVALTANSYKDKOKKCLSVGMNDFL 1950  
QY 118 AKPL-TDKIIPLI 130  
DB 1951 QPKIKTSLLIQMI 1964

RESULT 2

G65010 sensor protein evgs (EC 2.7.3.-) precursor - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
C:Accession: G65010; J00221; I41200  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.,  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: G65010  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1197 <BLAT>  
A:Cross-references: GB:AE000325; GB:U00096; NID:g1788709; PIDN:AACT5429.1; PID:g1788713  
A:Experimental source: strain K-12, substrain MG1655  
R:Utsumi, R.  
submitted to JIPID, January 1993  
A:Description: Newly identified genes involved in signal transduction of *Escherichia coli*  
A:Reference number: JU0220  
A:Accession: JU0221  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-151, 'F', 153-241, 'PL', 244-274, 'R', 276-419, 'FE', 422-738, 'D', 740-757, 'K', 759-  
R:Utsumi, R.; Katsuyama, S.; Taniguchi, M.; Horie, T.; Ikeda, M.; Igaki, S.; Nakagawa, H.  
Gene 140, 73-77, 1994  
A:Title: Newly identified genes involved in the signal transduction of *Escherichia coli*  
A:Reference number: I41198; MUID:94171083; PMID:8125343  
A:Accession: I41200  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-151, 'F', 153-241, 'PL', 244-274, 'R', 276-419, 'FE', 422-738, 'D', 740-757, 'K', 759-  
A:Cross-references: GB:D14008; NID:9456162; PIDN:BAA03108.1; PID:g216554  
C:Genetics:  
A:Gene: evgs  
C:Superfamily: evgs protein; response regulator homology  
C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase;  
F:687-935/Domain: sensor histidine kinase homology <SHK>  
F:661-1070/Domain: response regulator homology <RRH>  
F:721/Binding site: phosphate (His) (covalent) #status predicted  
F:1003/Binding site: phosphate (Asp) (covalent) #status predicted  
Query Match 26.3%; Score 178.5; DB 1; Length 1197;  
Best Local Similarity 34.4%; Pred. No. 5.1e-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;  
QY 12 IKKLNVLIVDDPLNLIIHEKIIKAIGISQTANNGEAAVIIHRDGGSSFDLILMDKEM 71  
DB 955 LPEKLSILIADDDHPTNRLLLKRLNLLGYDVEATDGVQA--LHKVSMQHYDLITDVM 1012  
QY 72 PERDGVSTTKKREMEVKSMIVGVTSADNEERAFMEAGLNHCLAKPLTKDKIPLIN 131  
DB 1013 PNVDGFELTRKLRQNSLPIWGLTANA-QANEREKGLCGMNLCLFKPLTLDVLKTHLS 1071  
QY 132 QL 133  
DB 1072 QL 1073  
RESULT 3  
A91035  
probable sensor for regulator EvgA [imported] - *Escherichia coli* (strain O157:H7, substrain  
C:Species: *Escherichia coli*  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 24-Aug-2001  
C:Accession: A91035  
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A91035  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1197 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA036672.1; PID:g13362719; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: Ecs3249  
C:Superfamily: evgs protein; response regulator homology  
Query Match 26.0%; Score 176.5; DB 2; Length 1197;  
Best Local Similarity 34.4%; Pred. No. 7.5e-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;  
QY 12 IKKLNVLIVDDPLNLIIHEKIIKAIGISQTANNGEAAVIIHRDGGSSFDLILMDKEM 71  
DB 955 LPEKLSILIADDDHPTNRLLLKRLNLLGYDVEATDGVQA--LHKVSMQHYDLITDVM 1012  
QY 72 PERDGVSTTKKREMEVKSMIVGVTSADNEERAFMEAGLNHCLAKPLTKDKIPLIN 131  
DB 1013 PNVDGFELTRKLRQNSLPIWGLTANA-QANEREKGLCGMNLCLFKPLTLDVLKTHLS 1071  
QY 132 QL 133  
DB 1072 QL 1073  
RESULT 5  
H87640  
sensor histidine kinase/response regulator [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87640  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: H87640  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <STO>  
A:Cross-references: GB:AE005673; NID:g13424832; PIDN:AAK25124.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3162  
Query Match 25.7%; Score 174; DB 2; Length 394;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;

QY 12 IKKLNVLIVDDPLNLIIHEKIIKAIGISQTANNGEAAVIIHRDGGSSFDLILMDKEM 71  
DB 955 LPEKLSILIADDDHPTNRLLLKRLNLLGYDVEATDGVQA--LHKVSMQHYDLITDVM 1012  
QY 72 PERDGVSTTKKREMEVKSMIVGVTSADNEERAFMEAGLNHCLAKPLTKDKIPLIN 131  
DB 1013 PNVDGFELTRKLRQNSLPIWGLTANA-QANEREKGLCGMNLCLFKPLTLDVLKTHLS 1071  
QY 132 QL 133  
DB 1072 QL 1073  
RESULT 4  
C85879  
probable sensor for regulator EvgA evgs [imported] - *Escherichia coli* (strain O157:H7  
C:Species: *Escherichia coli*  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: C85879  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: C85879  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1197 <STO>  
A:Cross-references: GB:AE005174; NID:g12516740; PIDN:AAG57495.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: evgs  
C:Superfamily: evgs protein; response regulator homology  
Query Match 26.0%; Score 176.5; DB 2; Length 1197;  
Best Local Similarity 34.4%; Pred. No. 7.5e-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;  
QY 12 IKKLNVLIVDDPLNLIIHEKIIKAIGISQTANNGEAAVIIHRDGGSSFDLILMDKEM 71  
DB 955 LPEKLSILIADDDHPTNRLLLKRLNLLGYDVEATDGVQA--LHKVSMQHYDLITDVM 1012  
QY 72 PERDGVSTTKKREMEVKSMIVGVTSADNEERAFMEAGLNHCLAKPLTKDKIPLIN 131  
DB 1013 PNVDGFELTRKLRQNSLPIWGLTANA-QANEREKGLCGMNLCLFKPLTLDVLKTHLS 1071  
QY 132 QL 133  
DB 1072 QL 1073



Db 869 RLTORIQLGLTLPIGVITANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 9

ES5861

Hypothetical protein rscC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: E85861

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: GB:AE005174; NID:g12516556; PIDN:AA57353.1; GSPDB:GN00145; UWGP:Z34

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: rscC

C:Superfamily: rscC protein; response regulator homology

Query Match 25.3%; Score 171.5; DB 2; Length 933;

Best Local Similarity 33.9%; Pred. No. 1.5e-06;

Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

Qy 18 LVIVDDPDLNLIHEKIIKAIGISQTSNNGEAVIHRDGGSSFDLILMDKEMPERDGV 77

Db 811 ILVDDHPINRLADQLTGLYRVVTANDGVDALNVL--SKNIDIVLSDVNPNDGY 868

Qy 78 STTKKLREMEYKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126

Db 869 RLTORIQLGLTLPIGVITANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 10

AD0149

two component sensor kinase/response regulator protein RscC (EC 2.7.3.-) [imported] - Yersinia pestis

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD0149

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Lil, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Cross-references: GB:AL590842; PIDN:CAC90055.1; PID:g15979276; GSPDB:GN00175

A:Accession: AD0149

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-957 <RUR>

A:Cross-references: GB:AL590842; PIDN:CAC90055.1; PID:g15979276; GSPDB:GN00175

C:Genetics:

A:Gene: rscC

C:Superfamily: rscC protein; response regulator homology

C:Keywords: phosphotransferase

Query Match 25.3%; Score 171.5; DB 2; Length 937;

Best Local Similarity 35.1%; Pred. No. 1.5e-06;

Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;

Qy 16 LNVIVDDPDLNLIHEKIIKAIGISQTSNNGEAVIHRDGGSSFDLILMDKEMPERD 75

Db 835 LQILVDDHPINRLADQLTGLYRVVTANDGVDALNVL--NTVDMVLTVNPNMD 892

Qy 76 GVSTTKKLREMEYKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126

Db 893 GYRLTERLQNLNHNFIIGVTANALAEKQRC-IEAGMDCNLSRPVTLDTL 942

RESULT 11

EB7644

sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: E87644

R:Rieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, N.J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

A:Title: Complete genome sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <STO>

A:Cross-references: GB:AE005673; NID:g13424865; PIDN:AAK25153.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3191

Query Match 24.0%; Score 162.5; DB 2; Length 642;

Best Local Similarity 33.0%; Pred. No. 5.5e-06;

Matches 36; Conservative 28; Mismatches 40; Indels 5; Gaps 2;

Qy 16 LNVIVDDPDLNLIHEKIIKAIGISQTSNNGEAVIHRDGGSSFDLILMDKEMPERD 75

Db 514 LHVLLVEDNATNRLIATRLMEALGARVTAEAGQGVAAARQG---FDLILMDIQMPVMD 570

Qy 76 GVSTTKKLREMEYKSMIVGVTSLADN--EERRAFMEAGLNHCLAKPLT 122

Db 571 GVEATHIRAFNSPAGAAPILAMTANAKAHQOASVLAAGMDGAIKAPLS 619

RESULT 12

AD0432

aerobic respiration control sensor/response regulatory protein (EC 2.7.3.-) [imported]

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD0432

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Lil, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Cross-references: GB:AL590842; PIDN:CAC92784.1; PID:g15981477; GSPDB:GN00175

A:Accession: AD0432

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92784.1; PID:g15981477; GSPDB:GN00175

C:Genetics:

A:Gene: arcB

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator homology

C:Keywords: phosphotransferase

Query Match 24.0%; Score 162.5; DB 2; Length 778;

Best Local Similarity 27.0%; Pred. No. 6.8e-06;

Matches 33; Conservative 38; Mismatches 48; Indels 3; Gaps 2;

Qy 16 LNVIVDDPDLNLIHEKIIKAIGISQTSNNGEAVIHRDGGSSFDLILMDKEMPERD 75

Db 525 LHVLLVEDIENLVIVARSVLEKLGNSVDVAMNGHDALAMFNP--EDFDVLDDIQLPDMS 582

Qy 76 GVSTTKKLREMEYKSMIVGVTSLADNE--EERRAFMEAGLNHCLAKPLTKDKI 134

Db 583 GLDIARQIRAEYCKQSILPLVALTANVLKDKKEYLDAGMDVDVLSKPLSVLPALTAMIKQFV 642

Qy 135 DA 136

Db 643 DS 644

RESULT 13

A69487

response regulator homolog - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Mar-2003  
C:Accession: A69487  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: A69487  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-117 <KLE>  
A:Cross-references: GB:AE000971; GB:AE000782; NID:92689294; PIDN:AA89351.1; PID:g264864  
C:Superfamily: signal transduction receiver (phosphoacceptor) protein (CheY-like); respo  
C:Keywords: phosphoprotein  
F:4-110/Domain: response regulator homology <RRH>  
F:50/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 23.5%; Score 159; DB 2; Length 117;  
Best Local Similarity 28.3%; Pred. No. 1.5e-06;  
Matches 36; Conservative 42; Mismatches 31; Indels 18; Gaps 5;

Qy 16 LNVLIYDD-----DPLNLIHE-KIKAIGGISQTANNGEAVIIHRDGGSSFDLILMDK 69  
Db 2 IKVWVDDTEAMREILKIMDKYVIE-----ASNGREAVEVLYRE-ERPDIIVLMDV 51

Qy 70 EMERQGVSTTKLRMEVKSMTIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIPL 129  
Db 52 MPMPLMGIEATSEIKIDPDAKVATVATVASSKGEK--VIEAGADYILKKPFTREKVEL 109

Qy 130 INQLMDA 136  
Db 110 IKRILNS 116

\*RESULT 14  
C82151 sensor histidine kinase VC1831 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: C82151  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.  
L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: C82151  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-736 <HEI>  
A:Cross-references: GB:AE004259; GB:AE003852; NID:9656353; PIDN:AAF94979.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1831  
A:Map position: 1

Query Match 23.5%; Score 159; DB 2; Length 736;  
Best Local Similarity 30.8%; Pred. No. 1.3e-05;  
Matches 41; Conservative 35; Mismatches 45; Indels 12; Gaps 5;

Qy 8 DIEKIKK-----LNVLIYDDPLNLIHEKIIKAIGGISQTANNGEAVIIHRDGGSSFD 63  
Db 590 DVKPTKATYPLGLRVLIVEDNRTNIMILEAFMRNKGFECHVMDGVQAITALQEE--SSFD 647

Qy 64 LILMDKEMPERDGVSTTKLRME--BVKSNIVGVTSLAD-NEERRAFMEAGLNELCLAK 119  
Db 648 LVLMDNMPLKDGIOATREINQLPQAKILLFGCT--ADVFKDRDKMLSGAGADDIAR 705

Qy 120 PLTKDKIIPLINQ 132  
Db 706 PIAEHEDLWALEQ 718

RESULT 15  
AG1897  
two-component hybrid sensor and regulator all0729 [imported] - Nostoc sp. (strain PCC  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AG1897  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AG1897  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1645 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA000019; PID:g17130074; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all0729

Query Match 23.5%; Score 159; DB 2; Length 1645;  
Best Local Similarity 27.7%; Pred. No. 3.1e-05;  
Matches 39; Conservative 40; Mismatches 50; Indels 12; Gaps 3;

Qy 3 TKSNGDIEKIKKLNLYIVDDPLNLIHEKIIKAIGGISQTANNGEAVIIHRDGGSSF 62  
Db 1414 TRSVIGLAPQRECRILVYDDVADSRLLVLLSVGFVQVQEAANGQALAIWQWHP-- 1471

Qy 63 DLILMDKEMPERDGVSTTKLRMEVK-----SMIVGVTSLADNEERRAFMEAGL 113  
Db 1472 QLILMDMRPIMDGYEATHFIRSAEIEHTTIPNPTIIIALTAFA-PEEQOAMLVQGC 1530

Qy 114 NHCLAKPLTKDKIIPLINQLM 134  
Db 1531 DDLINKPFSEKEILEKLNKYL 1551

Search completed: August 14, 2003, 16:51:17  
Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 23 Seconds  
(without alignments)  
278.071 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678

Sequence: 1 MATKSMGDIEIKIKLVLI.....IAKPLTKKIPLINQLMDA 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	178.5	26.3	1197	1 EVGS_ECOLI	P30855 escherichia
2	176.5	26.0	1197	1 EVGS_ECO57	P58402 escherichia
3	173.5	25.6	948	1 RCSC_SALTI	Q56128 salmonella
4	172.5	25.4	948	1 RCSC_SALTY	P58662 salmonella
5	171.5	25.3	949	1 RCSC_ECOLI	P14376 escherichia
6	156.5	23.1	777	1 FRZE_WXXA	P18769 myxococcus
7	150	22.1	622	1 SKN7_YEAST	P38989 saccharomyc
8	149.5	22.1	778	1 ARCB_ECO57	P58363 escherichia
9	149.5	22.1	778	1 ARCB_ECOLI	P22763 escherichia
10	145	21.4	1220	1 SLN1_YEAST	P39928 saccharomyc
11	143.5	21.2	712	1 SSK1_YEAST	Q07084 saccharomyc
12	142	20.9	907	1 GACS_PSESY	P48027 pseudomonas
13	140.5	20.7	226	1 DCTR_BACSU	P96602 bacillus su
14	140.5	20.7	230	1 DCTR_BACSD	Q9k998 bacillus ha
15	134	19.8	119	1 CHEY_ECO57	P24072 bacillus su
16	134	19.8	1238	1 BVGS_BORPE	P16575 bordetella
17	133.5	19.7	228	1 DCTR_BACME	P39486 bacillus me
18	133.5	19.7	918	1 BARA_ECOLI	P26607 escherichia
19	132.5	19.5	131	1 YSOL_PLEBO	P51586 plectonema
20	132.5	19.5	918	1 BARA_SHIFL	P59342 shigella fl
21	132	19.5	859	1 LUXO_VIBHA	P54302 vibrio harv
22	131.5	19.4	441	1 ZRAR_SALTI	Q82333 salmonella
23	131.5	19.4	441	1 ZRAR_SALTY	P25852 salmonella
24	130	19.2	1238	1 BVGS_BORPA	P40330 bordetella
25	128.5	19.0	849	1 LUXN_VIBHA	P54301 vibrio harv
26	127.5	18.8	522	1 MCS4_SCHPO	P87323 schizosacch
27	127	18.7	1238	1 BVGS_BORBR	P26762 bordetella
28	126.5	18.7	227	1 CPXR_BAEIN	P44895 haemophilus
29	125.5	18.5	443	1 ZRAR_KLEOX	Q9apd9 klebsiella
30	124	18.3	144	1 CHEY_TREPA	P96126 treponema p
31	124	18.3	539	1 PRR1_SCHPO	O14283 schizosacch
32	122.5	18.1	441	1 ZRAR_ECOLI	P14375 escherichia
33	122	18.0	248	1 ALGR_PSEAE	P26275 pseudomonas

#### RESULT 1

ID	EVGS_ECOLI	STANDARD	PRT	1197 AA
AC	P30855; P77644; Q9RF36; Q9RF37;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Sensor protein evgs precursor (EC 2.7.3.-).			
GN	EVGS OR B2370.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=94171083; PubMed=8125343;			
RA	Utsumi R., Katayama S., Taniguchi M., Horie T., Ikeda M., Igaki S.,			
RA	Nakagawa H., Miwa A., Tanabe H., Noda M.;			
RT	"Newly identified genes involved in the signal transduction of			
RT	Escherichia coli K-12.";			
RL	Gene 140:73-77(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=93173621; PubMed=1289796;			
RA	Utsumi R., Katayama S., Ikeda M., Igaki S., Nakagawa H., Miwa A.,			
RA	Taniguchi M., Noda M.;			
RT	"Cloning and sequence analysis of the evgAS genes involved in signal			
RT	transduction of Escherichia coli K-12.";			
RL	Nucleic Acids Symp. Ser. 27:149-150(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A. (VARIANTS EVGS1 AND EVGS4).			
RC	STRAIN=K12;			
RX	MEDLINE=20378313; PubMed=10923791;			
RA	Kato A., Ohnishi H., Yamamoto K., Furuta E., Tanabe H., Utsumi R.;			
RT	"Transcription of emrKY is regulated by the EvgA-Evgs two-component			
RT	system in Escherichia coli K-12.";			
RL	Biosci. Biotechnol. Biochem. 64:1203-1209(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474(1997).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=97349980; PubMed=9205837;			
RA	Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Tsono K.,			
RA	Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,			
RA	Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,			

Q8x613 escherichia  
P10957 escherichia  
Q34534 bacillus su  
Q45994 caulobacter  
P16244 escherichia  
Q44006 alcaligenes  
P49246 xanthomonas  
Q56312 thermotoga  
P31358 porphyra pu  
P43501 pseudomonas  
Q06065 escherichia  
O78428 guillardia

#### ALIGNMENTS

RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,  
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
RA Yamagata S., Horiuchi T.;  
RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*  
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and  
RT analysis of its sequence features.";  
RL DNA Res. 4:91-113(1997).  
RN [6]  
RP CHARACTERIZATION.  
RX MEDLINE=98194702; PubMed=9535079;  
RA Perraud A., L., Kimmel B., Weiss V., Gross R.;  
RT "Specificity of the evgAs and EvgAS phosphorelay is mediated by the  
RT C-terminal Hpt domains of the sensor proteins.";  
RA Mol. Microbiol. 27:875-887(1998).  
CC -!- FUNCTION: Member of the two-component regulatory system evgS/evgA.  
CC Phosphorylates evgA via a four-step phosphorelay in response to  
CC environmental signals.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Probable).  
CC -!- PM: Activation requires a sequential transfer of a phosphate  
CC group from a His in the primary transmitter domain, to an Asp in  
CC the receiver domain and to a His in the secondary transmitter  
CC domain.  
CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
CC -!- SIMILARITY: Contains 1 HPT domain.  
CC -!- SIMILARITY: Contains 1 response regulatory domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D14008; BAA03108.1; -;  
DR EMBL: AF201840; AAF17563.1; -;  
DR EMBL: AF201841; AAF17564.1; -;  
DR EMBL: AE000325; AAC75429.1; -;  
DR EMBL: D90867; BAA16241.1; -;  
DR PIR: G65010; G65010.  
DR HSP: P06143; IUDR.  
DR EcoGene: EG11610; evgs.  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003661; His\_kinA.  
DR InterPro: IPR005467; His\_kinase.  
DR InterPro: IPR002570; Hpt.  
DR InterPro: IPR001789; Response\_reg.  
DR InterPro: IPR001311; SBP/glu\_receptor.  
DR InterPro: IPR001638; SBP\_bac\_3.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; HiskA; 1.  
DR Pfam: PF00072; response\_reg; 1.  
DR Pfam: PF00497; SBP\_bac\_3; 1.  
DR PRINTS: PR00344; BCTRISENSOR.  
DR ProDom: PD000039; Response\_reg; 1.  
DR SMART: SM00387; HATPase\_c; 1.  
DR SMART: SM00388; HiskA; 1.  
DR SMART: SM00073; Hpt; 1.  
DR SMART: SM00052; PBPb; 2.  
DR SMART: SM00448; REC; 1.  
DR SMART: PS50109; HIS\_KIN; 1.  
DR PROSITE: PS0894; HPT; 1.  
DR PROSITE: PS50110; RESPONSE\_REGULATORY; 1.  
KW Sensory transduction; Transferrase; Kinase; Phosphorylation;  
KW Transmembrane; Inner membrane; Signal; Complete proteome.  
FT SIGNAL 1 21  
FT CHAIN 22 1197  
FT DOMAIN 22 325  
FT TRANSMEM 326 346  
FT DOMAIN 347 537  
FT TRANSMEM 538 558  
FT POTENTIAL.

FT DOMAIN 559 1197  
FT DOMAIN 718 938  
FT DOMAIN 960 1074  
FT DOMAIN 1098 1197  
FT MOD\_RES 721 721  
FT MOD\_RES 1009 1009  
FT MOD\_RES 1137 1137  
FT VARIANT 577 577  
FT VARIANT 701 701  
FT CONFLICT 152 152  
FT CONFLICT 242 243  
FT CONFLICT 275 275  
FT CONFLICT 420 421  
FT CONFLICT 739 739  
FT CONFLICT 758 758  
FT CONFLICT 761 761  
FT CONFLICT 877 877  
FT CONFLICT 1045 1045  
FT CONFLICT 1074 1074  
FT CONFLICT 1074 1074  
SQ SEQUENCE 1197 AA; 134742 MW; E8E1DE0F797B1278 CRC64;  
Query Match 26.3%; Score 178.5; DB 1; Length 1197;  
Best Local Similarity 34.4%; Pred No. 2.9e-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;  
QY 12 IKKLVNLIIVDDPLNLIHEKIKAIIGISQQTANNGEEAVIIHRDGGSPFDLIMDKEM 71  
Db 955 LPEKLSIIADDPHTNRLKLLKQLNLLGYDDEATDGVQA--LHKVSNQHVDDLLITDVM 1012  
QY 72 PERDGVSTTKKREMEVKSMIVGTSLADNEEERAFMEAGLNHCLAKPLTKDKIIPLN 131  
Db 1013 PNMDGFEIIRKREONSSIPWGLTANA-QANEREKGLSCGMNLCFLPDLVLTLS 1071  
QY 132 QL 133  
Db 1072 QL 1073  
RESULT 2  
EVGS\_ECO57 STANDARD; PRT: 1197 AA.  
ID EVGS\_ECO57  
AC P38402;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Sensor protein evgs precursor (EC 2.7.3.-).  
GN EVGS OR 23632 OR ECS3249.  
OS *Escherichia coli* O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; *Escherichia*.  
OC NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
RL Nature 409:529-533(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";

	Query Match	26.08;	Score 176.5;	DB 1;	Length 1197;
	Best local Similarity	34.4%;	Pred.No. 4.3e-07;		
	Matches	42;	Conservative	25;	Mismatches 52; Indels 3; Gaps 2;
QY	12	IKKXINLVIVDDPPLNLIHKEIKAIIGGISGTANNGEAAVIIHRDGGSSFDLIMDKEM	71		
Dd	95S	LPEKLISLLADHPTEPNLLKKROLNLGVDVEADTGVOA--LHKVSMDHYDLITIVNM	1012		

QY 12 IKKKLVLIIVDDPLNIIHEKIIKAIGGISOTANNGEEAVIHRDGGSSFDILMDKEM 71  
:  
Db 955 LPEKLSIIADDPHTNBLLLRNLGLGYDVDEATDCGQA--LHKVSMQHVDLLITBVNM 1012

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR EMBL: AL627274; CAD07502.1; -.
DR EMBL: AE016836; AAC69299.1; -.
DR EMBL: X87830; CAA61095.1; -.
DR HSSP: P06143; IUDR.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinase.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR PRINTS: PR00072; response_reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50112; PAS; FALSE_NEG.
DR Sensory transduction; Transferrase; Kinase; Bacterial capsule;
KW Inner membrane; Transmembrane; Phosphorylation; Complete proteome.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 POTENTIAL.
FT DOMAIN 42 313 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 314 334 POTENTIAL.
FT DOMAIN 335 948 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 357 425 PAS.
FT DOMAIN 476 692 HISTIDINE KINASE.
FT DOMAIN 826 940 RESPONSE REGULATORY.
FT MOD_RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 875 875 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 948 AA; 106237 MW; AE3A21701265A865 CRC64;
Query Match 25.6%; Score 173.5; DB 1; Length 948;
Best Local Similarity 33.9%; Pred. No. 5.9e-07;
Matches 37; Conservative 33; Mismatches 36; Indels 3; Gaps 2;
QY 18 VLIYDDPLNLIIEKIKAIGISQTANNGEAVIHRDGGSSFDLILMDKEMPERDGV 77
D5 827 ILVVDHPINRLLADQLGSLGYQCKTANDGVDAIIVL--SKNAIDIVLSVNNPMDGY 884
QY 78 STTKKLEMEVKSMIVGVTSIADNEERAFMEAGLNHCLAKPLTKDXI 126
D5 885 RLTORIRQLGLTLFVVGVTANALAEKQRC-LESGMDSCLSKPVTLDAL 932

```

## RESULT 4

```

RCSC_SALTY STANDARD; PRT; 948 AA.
AC PS9662.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sensor protein rscC (EC 2.7.3.-) (Capsular synthesis regulator
DE component C).
DE RCSC OR STM2271.
GN Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

```

```

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Member of the two-component regulatory system rscC/rscB
CC involved in the regulation of the expression of genes involved in
CC colanic acid capsule synthesis. RscC probably functions as a
CC membrane-associated protein kinase that phosphorylates rscB in
CC response to environmental signals.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PH: Activation probably requires a transfer of a phosphate group
CC between a His in the transmitter domain and an Asp of the receiver
CC domain (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).


DR EMBL: AE008801; AAL21172.1; -.
DR StyGene: SG2727; rscC.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinase.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50112; PAS; FALSE_NEG.
DR Sensory transduction; Transferrase; Kinase; Bacterial capsule;
KW Inner membrane; Transmembrane; Phosphorylation; Complete proteome.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 POTENTIAL.
FT DOMAIN 42 313 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 314 334 POTENTIAL.
FT DOMAIN 335 948 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 357 425 PAS.
FT DOMAIN 476 692 HISTIDINE KINASE.
FT DOMAIN 826 940 RESPONSE REGULATORY.
FT MOD_RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 875 875 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 948 AA; 106279 MW; BAAD8DA557D5868B CRC64;
Query Match 25.4%; Score 172.5; DB 1; Length 948;
Best Local Similarity 34.6%; Pred. No. 7.1e-07;
Matches 37; Conservative 32; Mismatches 35; Indels 3; Gaps 2;
QY 18 VLIYDDPLNLIIEKIKAIGISQTANNGEAVIHRDGGSSFDLILMDKEMPERDGV 77
D5 827 ILVVDHPINRLLADQLGSLGYQCKTANDGVDAIIVL--SKNAIDIVLSVNNPMDGY 884
QY 78 STTKKLEMEVKSMIVGVTSIADNEERAFMEAGLNHCLAKPLTKD 124
D5 885 RLTORIRQLGLTLFVVGVTANALAEKQRC-LESGMDSCLSKPVTLD 930


```



01-NOV-1990 (Rel. 16, Last sequence update)  
 28-FEB-2003 (Rel. 41, Last annotation update)  
 Gliding motility regulatory protein (EC 2.7.3.-).  
 FRZE  
 GN Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;  
 OC Cytophacterales; Myxococcaceae; Myxococcus.  
 OC NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90332690; PubMed=2165608;  
 RA McCleary W.R., Zusman D.R.;  
 RT "Frze of Myxococcus xanthus is homologous to both Chea and Chey of  
 RL Salmorella typhimurium".  
 Proc. Natl. Acad. Sci. U.S.A. 87:5898-5902(1990).  
 RN [2]  
 RP PHOSPHORYLATION OF HIS-49.  
 RX MEDLINE=91072208; PubMed=2123853;  
 RA McCleary W.R., Zusman D.R.;  
 RT "Purification and characterization of the Myxococcus xanthus Frze  
 RL protein shows that it has autophosphorylation activity".  
 J. Bacteriol. 172:6661-6668(1990).  
 CC -!- FUNCTION: FRZE IS INVOLVED IN A SENSORY TRANSDUCTION PATHWAY THAT  
 CC CONTROLS THE FREQUENCY AT WHICH CELLS REVERSE THEIR GLIDING  
 CC DIRECTION. FRZE SEEMS TO BE CAPABLE OF AUTOPHOSPHORYLATING ITSELF  
 CC ON AN HISTIDINE RESIDUE AND THEN TO TRANSFER THAT GROUP TO AN  
 CC ASPARTATE RESIDUE IN THE C-TERMINAL PART OF THE PROTEIN.  
 CC -!- SIMILARITY: Contains 1 cheW-like domain.  
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
 CC -!- SIMILARITY: Contains 1 HPT domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M35192; AAA25396.1; -;  
 DR PIR; A35966; A35966.  
 DR HSP; Q56310; LB3Q.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR002545; CheW.  
 DR InterPro; IPR003467; His\_kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF01584; CheW; 1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF01627; Hpt; 1.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR PRINTS; PR00344; BCTRLSENSOR.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00260; CheW; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS0851; CHEW; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 DR PROSITE; PS50894; HPT; 1.  
 DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
 KW Sensory transduction; Transference; Kinase; Phosphorylation.  
 FT DOMAIN 1 108 HPT.  
 FT DOMAIN 270 509 HISTIDINE KINASE.  
 FT DOMAIN 511 645 CHEW-LIKE.  
 FT DOMAIN 660 776 RESPONSE REGULATORY.  
 FT MOD\_RES 49 49 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT DOMAIN 130 197 ALA/PRO-RICH (POSSIBLE HINGE REGION).  
 SQ SEQUENCE 777 AA; 83189 MW; 9912BD40991C59E5 CRC64;  
 Query Match 23.1%; Score 156.5; DB 1; Length 777;

Best Local Similarity 35.2%; Pred. No. 1.2e-05;  
 Matches 38; Conservative 27; Mismatches 38; Indels 5; Gaps 3;  
 Qy 14 KKLNVLYDDPLNLIHKKIIRKIGISQTANGGEAVIIHRDGGSFLLMDKEMPE 73  
 Db 657 KRLRVLLVDDSPITARATGALVGHVVEAQDGEAYV--KVQNTYDLILTDVQMPK 714  
 Qy 74 RDGVSTTKKLEMEVKSMI--VGVTSLADNEEERRAFMEAGLNHCLAK 119  
 Db 715 LDGFSLARLRLKSTPAVARIPVILLSSLAGSPEDKRRG-LDAGADAYLVK 761

RESULT 7  
 SKN7\_YEAST STANDARD; PRT; 622 AA.  
 AC P38869; P39747;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative transcription factor SKN7 (POS9 protein).  
 GN SKN7 OR POS9 OR BRV1 OR YHR206W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94042854; PubMed=8226633;  
 RA Brown J.L., North S., Bussey H.;  
 RT "SKN7, a yeast multicopy suppressor of a mutation affecting cell wall  
 RT beta-glucan assembly, encodes a product with domains homologous to  
 RT prokaryotic two-component regulators and to heat shock transcription  
 RT factors".  
 RL J. Bacteriol. 175:6908-6915(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Krens B., Charizanis C., Entian K.-D.;  
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 RA Du Z., Favell A., Fulton L., Gattling S., Geisel C., Kirsten J.,  
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,  
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,  
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,  
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaudin M.;  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII".  
 RL Science 265:2077-2082(1994).  
 RN [4]  
 RP FUNCTION, AND MUTAGENESIS.  
 RX MEDLINE=95045411; PubMed=7957083;  
 RA Brown J.L., Bussey H., Stewart R.C.;  
 RT "yeast Skn7p functions in a eukaryotic two-component regulatory  
 RT pathway".  
 RL EMBO J. 13:5186-5194(1994).  
 CC -!- FUNCTION: INVOLVED IN OXIDATIVE STRESS. TRANSCRIPTION FACTOR THAT  
 CC MAY FUNCTION IN A TWO-COMPONENT SIGNAL TRANSDUCTION PATHWAY THAT  
 CC ACTS IN PARALLEL WITH THE PKC1 CASCADE TO REGULATE GROWTH AT THE  
 CC CELL SURFACE.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

```

CC EMBL; U00485; AAC48911.1; -
DR EMBL; X83031; CAA58143.1; -
DR EMBL; U00029; AAB69734.1; -
DR PIR; A49344; A49344.
DR HSP; P22121; 2HIS.
DR TRANSFAC; T03481; -.
DR SGD; S0001249; SKN7.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0003700; F:transcription factor activity; IDA.
DR GO; GO:000156; F:two-component response regulator activity; IDA.
DR GO; GO:0006350; P:transcription; IDA.
DR InterPro; IPR000232; HSF_DNA_bind.
DR InterPro; IPR002341; HSF_EIS.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00447; HSF_DNA_bind; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00056; HSFDOMAIN.
DR PRODOM; PD001788; HSF_DNA_bind; 1.
DR PRODOM; PD000039; Response_reg; 1.
DR SMART; SM00415; HSF; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00434; HSF_DOMAIN; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR Transcription regulation; Sensory transduction; Nuclear protein;
KW DNA-binding; Phosphorylation.
FT DNA_BIND 86 190
FT DOMAIN 378 492
FT MOD_RES 427 427
FT MUTAGEN 427 427
FT MUTAGEN 427 427
FT MUTAGEN 427 427
SQ SEQUENCE 622 AA; 69202 MW; 4C732FD66E326742 CRC64;

Query Match 22.1%; Score 150; DB 1; Length 622;
Best Local Similarity 28.1%; Pred. No. 3.3e-05;
Matches 38; Conservative 32; Mismatches 45; Indels 20; Gaps 4;

QY 12 IKKLNVLIVDDPLNLHIEKIIKAIGISQTANNGEAVII---HRDGGSSFDLLMD 68
-DB 373 LRKGHVLIVDDVAVISQLCSKFLRYGCTVQVSDGLSAISTLEKRY-----YDLVMD 427
QY 69 KEMPERDGVSTKKIREMVKMSIVGVTSADNEERAFNEAGLNHCLAKPLTKDKI-- 126
-DB 428 IYVPLNDGATASIVRSFDENPPIATMGNIMN-QDLITYLQHGNVDILAKPFRDRLHS 486.
QY 127 -----IPLINQ 132
-DB 487 ILIRYLKDRIPLCEQ 501

RESULT 8
ID ARCB_ECO57
AC ARCB_ECO57 STANDARD; PRT; 778 AA.
AC P36363;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aerobic respiration control sensor protein arcB (EC 2.7.3.-).
GN ARCB OR 24574 OR ECS4089.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.,

```

```

"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
Nature 409:529-533(2001).
[2]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Member of the two-component regulatory system arcB/arcA.
CC Sensor-regulator protein for anaerobic repression of the arc
CC regulon. Activates arcA via a four-step phosphorylation. ArcB can
CC also dephosphorylate arcA by a reverse phosphorylation involving His-
CC 717 and Asp-576 (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 HPT domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF005549; AAG88344.1; -.
CC EMBL; AF002564; BAB37512.1; -.
CC FIR; A91140; A91140.
CC FIR; D85985; D85985.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR003661; His_kinA.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR002570; Hpt.
CC InterPro; IPR000700; PAS-assoC_C.
CC InterPro; IPR000014; PAS_domain.
CC InterPro; IPR001789; Response_reg.
CC Pfam; PF02518; HATPase_C; 1.
CC Pfam; PF00512; HSKA; 1.
CC Pfam; PF00989; PAS; 1.
CC Pfam; PF00072; response_reg; 1.
CC ProDom; PD000039; Response_reg; 1.
CC SMART; SM00387; HATPase_C; 1.
CC SMART; SM00388; HSKA; 1.
CC SMART; SM00073; HPT; 1.
CC SMART; SM00091; PAS; 1.
CC SMART; SM00448; REC; 1.
CC SMART; SM00448; REC; 1.
CC TIGRfams; TIGR00229; sensory_box; 1.
CC PROSITE; PS0109; HIS_KIN; 1.
CC PROSITE; PS0894; HPT; 1.
CC PROSITE; PS0113; PAC; 1.
CC PROSITE; PS0112; PAS; 1.
CC PROSITE; PS0110; RESPONSE_REGULATORY; 1.
CC Sensory transduction; Transferase; Kinase; Phosphorylation;
CC Transmembrane; Inner membrane; Transcription regulation;
CC Complete proteome.
KW DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 26 46 POTENTIAL.
KW TRANSMEM 47 57 PERIPLASMIC (POTENTIAL).
KW TRANSMEM 58 78 POTENTIAL.
KW DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 153 223 PAS.
FT DOMAIN 226 278 PAC.
FT DOMAIN 289 507 HISTIDINE KINASE.
FT DOMAIN 527 643 RESPONSE REGULATORY.
FT DOMAIN 678 771 HPT.
FT MOD_RES 292 292 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 292 292 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 576 576 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 717 717 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 778 AA; 88010 MW; C8AE004B007F9D30 CRC64;

Query Match 22.1%; Score 149.5; DB 1; Length 778;
Best local Similarity 25.6%; Pred. No. 4.6e-05;
Matches 31; Conservative 40; Mismatches 47; Indels 3; Gaps 2;

QY 16 LNVLLVDDPDLNLIHKKIKAIGGISTQANNGEEAVIIHRDGGSSFDLILMDKEMPERD 75
LNVLLVDDPDLNLIHKKIKAIGGISTQANNGEEAVIIHRDGGSSFDLILMDKEMPERD 75
DB 526 LNVLLVDDPDLNLIHKKIKAIGGISTQANNGEEAVIIHRDGGSSFDLILMDKEMPERD 583
LNVLLVDDPDLNLIHKKIKAIGGISTQANNGEEAVIIHRDGGSSFDLILMDKEMPERD 583
QY 76 GVSTTKKLENEKSMIVGTVSLADNE-EERRAFMEAGLNHCLAKPLTKDKIIPLNQLM 134
GVSTTKKLENEKSMIVGTVSLADNE-EERRAFMEAGLNHCLAKPLTKDKIIPLNQLM 134
DB 584 GLDISREITKRYPREDLPLLVANLVKDKQEVNAGMDVLSKPLSVFALTMIKKFW 643
GLDISREITKRYPREDLPLLVANLVKDKQEVNAGMDVLSKPLSVFALTMIKKFW 643
QY 135 D 135
DB 644 D 644

RESULT 9
ARCB_ECOLI
ID ARCB_ECOLI STANDARD; PRT: 778 AA.
AC P22763;
DT 01-AUG-1991 (Rel. 19, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aerobic respiration control sensor protein arCB (EC 2.7.3.-).
GN ARCB OR B3210 OR SF3250.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 523;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=90355832; PubMed=2201868;
RA Tuchi S., Matsuda Z., Fujiwara T., Lin E.C.C.;
RT "The arcB gene of Escherichia coli encodes a sensor-regulator protein
RT for anaerobic repression of the arc modulon.";
RL Mol. Microbiol. 4:715-727(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN SEQUENCE FROM N.A.
RC SPECIES=S. flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [4]
RN CHARACTERIZATION.

```

```

RC SPECIES=E.coli; STRAIN=MJ5;
RX MEDLINE=97431492; PubMed=9286997;
RA Georgellis D., Lynch A.S., Lin E.C.C.;
RT "In vitro phosphorylation study of the arc two-component signal
RT transduction system of Escherichia coli.";
RL J. Bacteriol. 179:5429-5435(1997).
RN [5]
RN CHARACTERIZATION.
RC SPECIES=E.coli; STRAIN=MJ5;
RX MEDLINE=99047671; PubMed=9830034;
RA Georgellis D., Kwon O., De Wulf P., Lin E.C.C.;
RT "Signal decay through a reverse phosphorelay in the arc two-component
RT signal transduction system.";
RL J. Biol. Chem. 273:32864-32869(1998).
RN [6]
RN MUTAGENESIS OF HIS-292; ASP-576 AND HIS-717.
RC SPECIES=E.coli; STRAIN=K12 / MC4100;
RX MEDLINE=20309722; PubMed=10851007;
RA Kwon O., Georgellis D., Lin E.C.C.;
RT "Phosphorelay as the sole physiological route of signal transmission
RT by the arc two-component system of Escherichia coli.";
RL J. Bacteriol. 182:3858-3862(2000).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS) OF 660-778.
RC SPECIES=E.coli;
RX MEDLINE=97207018; PubMed=9054511;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Insights into multistep phosphorelay from the crystal structure of
RT the C-terminal HPT domain of ArcB.";
RL Cell 88:717-723(1997).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 659-776 IN COMPLEX WITH CHEY.
RC SPECIES=E.coli;
RX MEDLINE=98437504; PubMed=9761838;
RA Kato M., Mizuno T., Hakoshima T.;
RT "Crystallization of a complex between a novel C-terminal transmitter,
RT HPT domain, of the anaerobic sensor kinase ArcB and the chemotaxis
RT response regulator CheY.";
RL Acta Crystallogr. D 54:140-142(1998).
RN [9]
RN X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 659-776.
RC SPECIES=E.coli;
RX MEDLINE=2003135; PubMed=10531481;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Refined structure of the histidine-containing-phosphotransfer (HPT)
RT domain of the anaerobic sensor kinase ArcB from Escherichia coli at
RT 1.57-A resolution.";
RL Acta Crystallogr. D 55:1842-1849(1999).
RN [10]
RN FUNCTION: Member of the two-component regulatory system arcB/arcA.
CC Sensor-regulator protein for anaerobic repression of the arc
CC modulon. Activates arcA via a four-step phosphorelay. ArcB can
CC also dephosphorylate arcA by a reverse phosphorelay involving His-
CC 717 and Asp-576.
CC -! SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -! PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain.
CC -! SIMILARITY: Contains 1 histidine kinase domain.
CC -! SIMILARITY: Contains 1 HPT domain.
CC -! SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -! SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -! SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

EMBL; X53315; CAA37397.1; -  
 DR EMBL; U18997; AAS8012.1; -  
 DR EMBL; A500400; AAC76242.1; -  
 DR EMBL; A5015336; AAN44715.1; -  
 DR PIR; D65114; RGEAR.  
 DR PDB; 1A0B; 18-MAR-98.  
 DR PDB; 2A0B; 17-JUN-98.  
 DR PDB; 1BDJ; 11-MAY-99.  
 DR PDB; 1FR0; 31-DEC-02.  
 DR EcoGene; Egl10062; arcB.  
 DR InterPro; IPR003594; ATPbind\_Atpase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003681; His\_kinase.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR000700; PAS-assoc\_C.  
 DR InterPro; IPR000014; PAS\_domain.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00512; HSKA; 1.  
 DR Pfam; PF00989; PAS; 1.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR PRINTS; PR00344; BCTRLSENSOR.  
 DR PRODOM; PD000039; Response\_reg; 1.  
 DR SMART; SM00388; HSKA; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00091; PAS; 1.  
 DR SMART; SM00448; REC; 1.  
 DR TIGRFAMS; TIGR00229; sensory\_box; 1.  
 DR PROSITE; PS0109; HIS\_KIN; 1.  
 DR PROSITE; PS00894; HPT; 1.  
 DR PROSITE; PS0113; PAC; 1.  
 DR PROSITE; PS0112; PAS; 1.  
 DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
 DR Sensory\_transduction; Transferrase; Kinase; Phosphorylation;  
 DR Transmembrane; Inner membrane; Transcription regulation; 3D-structure;  
 DR Complete proteome.  
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 26 46 POTENTIAL.  
 FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 POTENTIAL.  
 FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 153 223 PAS.  
 FT DOMAIN 226 278 PAC.  
 FT DOMAIN 289 507 HISTIDINE KINASE.  
 FT DOMAIN 527 643 RESPONSE REGULATORY.  
 FT DOMAIN 678 771 HPT.  
 FT MOD\_RES 292 292 PHOSPHORYLATION (AUTO-).  
 FT MOD\_RES 576 576 PHOSPHORYLATION (PROBABLE).  
 FT MOD\_RES 717 717 PHOSPHORYLATION (PROBABLE).  
 FT MUTAGEN 292 292 H->Q: LOSS OF ACTIVITY.  
 FT MUTAGEN 576 576 D->A: LOSS OF ACTIVITY.  
 FT MUTAGEN 717 717 H->Q: LOSS OF ACTIVITY.  
 FT CONFLICT 469 470 MISSING (IN REF. 2).  
 SQ SEQUENCE 778 AA; 87982 MW; DD61EA6ECF95AD30 CRC64;  
 Query Match 22.18; Score 149.5; DB 1; Length 778;  
 Best Local Similarity 25.64; Pred. No. 4.6e-05;  
 Matches 31; Conservative 40; Mismatches 47; Indels 3; Gaps 2;  
 QY 16 LNVLIVDDPLNLIIHEKIIKAIGISQTANNGEAVIHRDGGSGFDLILDKEMPERD 75  
 DB 526 LNVLLVEDIENVIVARSVLEKLSNVDVAMTKGAALMEKPG-EYDLVLLDIQLPDMT 583  
 QY 76 GVSTTKKUREMVEKSMIVGVTSGLADNE-EERRAFMEAGLNHCLAKPLTKDKIPLINQLM 134  
 DB 584 GLDISRELTKYPRDELPLPLVALTANVLDKQVEYLNAGMDVLSKPLSPALTAIATIKFEW 643  
 QY 135 D 135  
 DB 644 D 644

RESULT 10  
 SLN1\_YEAST STANDARD; PRT; 1220 AA.  
 ID SLN1\_YEAST  
 AC P39928;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Osmolarity two-component system protein SLN1 (EC 2.7.3.-).  
 GN SLN1 OR YPD2 OR YII147C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / YPH1;  
 RX MEDLINE=94024010; PubMed=8211183;  
 RT Ota I.M., Varshavsky A.;  
 RL "A yeast protein similar to bacterial two-component regulators.";  
 RL Science 262:566-569(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / AB972;  
 RX PubMed=9169870;  
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,  
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,  
 RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,  
 RA Moulle S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,  
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";  
 RL Nature 387:84-87(1997).  
 RN [3]  
 RP MUTAGENESIS OF HIS-576 AND ASP-1144.  
 RX MEDLINE=94239498; PubMed=8183345;  
 RA Maeda T., Wurgler-Murphy S.M., Saito H.;  
 RT "A two-component system that regulates an osmosensing MAP kinase  
 cascade in yeast.";  
 RL Nature 369:242-245(1994).  
 CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM  
 SLN1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE  
 EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->  
 PBS2->HOG1 PATHWAY. INACTIVE SLN1 ALLOWS THE UNPHOSPHORYLATED SSK1  
 PROTEIN TO ACTIVATE SSK2 AND SSK22, TWO MAPKKS THAT FURTHER  
 STIMULATE THE PBS2-HOG1 MAPK CASCADE. IN LOW OSMOLARITY MEDIA,  
 THE ACTIVED SLN1 HISTIDINE KINASE REPRESSSES THE ACTIVATION OF THE  
 PBS2-HOG1 KINASE CASCADE THROUGH PHOSPHORYLATION OF SSK1.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- PTM: ACTIVATION PROBABLY REQUIRES A TRANSFER OF A PHOSPHATE GROUP  
 BETWEEN A HIS IN THE TRANSMITTER DOMAIN AND AN ASP OF THE RECEIVER  
 DOMAIN.  
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; Z38059; CAA86131.1; -  
 DR EMBL; U01835; AAC48912.1; -  
 DR PIR; S48387; S48387.  
 DR HSSP; P06143; 1UDR.  
 DR SGD; S0001409; SLN1.  
 DR GO; GO:0007234; P:osmosensory signaling pathway via two-compo. . . ; IDA.  
 DR GO; GO:0042542; P:response to hydrogen peroxide; IMP.  
 DR InterPro; IPR003594; ATPbind\_Atpase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003661; His\_kinase.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR001789; Response\_reg.

RESULT 12
GACS_PSES
ID GACS
AC P480
DT 01-F
DT 01-F

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sensor protein gacS (EC 2.7.3.-).  
 GN GACS OR LENA.  
 OS Pseudomonas syringae (pv. syringae).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 CC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=321;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92234961; PubMed=1314807;  
 RA Hrabak E.M., Willis D.K.;  
 RT "The lna gene required for pathogenicity of Pseudomonas syringae pv.  
 syringae on bean is a member of a family of two-component  
 regulators."  
 RL J. Bacteriol. 174:3011-3020(1992).  
 CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM  
 GACA/GACS(LENA). MAY BE INVOLVED IN LESION FORMATION, SWARMING AND  
 IN THE PRODUCTION OF EXTRACELLULAR PROTEASE, SYRINGOMYCIN AND N-  
 ACYL-L-HOMOSERINE LACTONE (ACYL-HSL). REQUIRED FOR PATHOGENICITY  
 ON BEAN.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 (Probable).  
 CC -!- PTM: Activation requires a sequential transfer of a phosphate  
 group from a His in the primary transmitter domain, to an Asp in  
 the receiver domain and to a His in the secondary transmitter  
 domain (By similarity).  
 CC -!- SIMILARITY: Contains 1 HAMP domain.  
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
 CC -!- SIMILARITY: Contains 1 HPT domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M80477; AAA25877.1; --  
 DR HSPP; P06143; IUDR.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003660; HAMP.  
 DR InterPro; IPR003661; His\_kinA.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF00672; HAMP; 1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00512; Hiska; 1.  
 DR Pfam; PF01627; Hpt; 1.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR PRINTS; PR00344; BCTRLENSOR.  
 DR PRODOM; PD000039; Response\_reg; 1.  
 DR SMART; SM00304; HAMP; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00388; Hiska; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS50885; HAMP; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 DR PROSITE; PS08994; HPT; 1.  
 DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
 KW Sensory transduction; Transpherase; Kinase; Phosphorylation;  
 KW Transmembrane; Inner membrane.  
 FT TRANSMEM 9 25  
 FT TRANSMEM 84 101 POTENTIAL.  
 FT TRANSMEM 159 178 POTENTIAL.  
 FT DOMAIN 182 234 HAMP.  
 FT DOMAIN 281 502 HISTIDINE KINASE.  
 FT DOMAIN 658 777 RESPONSE REGULATORY.  
 FT DOMAIN 814 907 HPT.

FT MOD\_RES 284 284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 707 707 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 853 853 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 907 AA; 99195 MW; 5B9F4663DAF3492C CRC64;  
 Query Match 20.9%; Score 142; DB 1; Length 907;  
 Best Local Similarity 27.6%; Pred. No. 0.00023;  
 Matches 35; Conservative 33; Mismatches 33; Indels 26; Gaps 4;  
 QY 18 VLIVDDPLNLIIEKIKKIGGISTQANNGEAAVIRHGG-----SFDLILMD 68  
 DB 659 VLVVDNPNALLVQTLLEDMG-----AEVVAVEGGVAAVNAVQQAFLVMD 707  
 QY 69 KEMPERDGVSTTKKLENEVKS-----MIVGVTSIADNEERERAEAGNHCCLKPLTK 123  
 DB 708 VQPMGDGRQATEAIRAWEARNOSSLPVALTAHA-MANEKRSLLQSGMDYLTKEPSE 766  
 QY 124 DKRIIPLI 130  
 DB 767 RLQAVV 773  
 RESULT 13  
 DCTR\_BACSU STANDARD; PRT: 226 AA.  
 ID DCTR\_BACSU AC P96602; P94503;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable C4-dicarboxylate response regulator dctr.  
 GN DCTR.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kasahara Y., Nakai S., Lee S., Sadaie Y., Ogasawara N.;  
 RT "A 148 kbp sequence of the region between 35 and 47 degree of the  
 Bacillus subtilis genome."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 Borriess R., Boursier L., Brans A., Braun M., Brignall S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,  
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 Kurita K., Lapius A., Lardinols S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Maestl D., Nakai S., Noback M.,  
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nagai S., Park S.H.,  
 Noore D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 Prasecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,  
 Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,  
 Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 Viari A., Wambutt R., Wedler E., Wedler K., Weitzenecker T.,  
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

"The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"  
 Nature 390:249-256(1997).  
 [3]  
 SEQUENCE OF 14-101 FROM N.A.  
 STRAIN=9202; JCM 9153;  
 MEDLINE=97311990; PubMed=9168601;  
 RA Morel-Deville F., Ehrlich S.D., Morel P.;  
 RT "Identification by PCR of genes encoding multiple response  
 regulators,"  
 RL Microbiology 143:1513-1520(1997).  
 [4]  
 FUNCTION, AND GENE NAME.  
 RT STRAIN=168; PubMed=10708364;  
 RX Asai K., Baik S.-H., Kasahara Y., Moriya S., Ogasawara N.;  
 RA "Regulation of the transport system for C4-dicarboxylic acids in  
 RT *Bacillus subtilis*,"  
 RL Microbiology 146:263-271(2000).  
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM DCTS/DCTR.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- PTM: PHOSPHORYLATED BY DCTS (PROBABLE).  
 CC -1- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AB001488; BAB19283.1;  
 DR EMBL: 299106; CAB12253.1;  
 DR EMBL: 082580; AAB41751.1;  
 DR PIR: B69771; B69771.  
 DR HSP: Q56312; ITWY.  
 DR Subtilist; BG12074; dctr.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF00072; response\_reg.1.  
 DR PIRSF: PIRSF006171; RR\_citrat\_malat; 1.  
 DR ProDom: PD000039; Response\_reg.1.  
 DR SMART: SM00448; REC.1.  
 DR HSP: Q56312; ITWY.  
 DR Subtilist; BG12074; dctr.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF00072; response\_reg.1.  
 DR PIRSF: PIRSF006171; RR\_citrat\_malat; 1.  
 DR ProDom: PD000039; Response\_reg.1.  
 DR SMART: SM00448; REC.1.  
 DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.  
 DR Sensory transduction; Transcription regulation; DNA-binding;  
 KW Activator; Phosphorylation; Complete proteome.  
 FT DOMAIN 7 123 RESPONSE REGULATORY.  
 FT MOD\_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).  
 FT DNA\_BIND 176 206 H-T-H MOTIF (POTENTIAL).  
 FT CONFLICT 24 24 F -> I (IN REF. 3).  
 SQ SEQUENCE 226 AA; 25539 MW; 018115B14E9EF47D CRC64;  
 Query Match 20.7%; Score 140.5; DB 1; Length 226;  
 Best Local Similarity 25.9%; Pred. No. 6.7e-05;  
 Matches 30; Conservative 39; Mismatches 42; Indels 5; Gaps 3;  
 QY 13 KXKLVVDDPLNLIHEKIIKAIGIS--QTANNGEAAVHHRDGGSSFDLIMDKKE 70  
 Db 3 RREKVLIEDDPVQVFNKQFIITVKGVTCATAGNCEGKLIKE--EQPDVLVDVY 60  
 QY 71 MPEDGVSTTKKREMEVKSIVGTSLADNEERAFMEAGLNHCLAKPLTKDKI 126  
 Db 61 MPKKGDIKTQIRKQKLEVDVI-VVSAKDKETISLMLONGAVDYILKPKFLERM 115  
 RESULT 14  
 DCTR\_BACHD STANDARD; PRT: 230 AA.  
 AC Q9K988;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable C4-dicarboxylate response regulator dctr.  
 GN DCTR OR BH2751.  
 OS *Bacillus halodurans*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C-125 / JCM 9153;  
 RC MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*  
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*,"  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM DCTS/DCTR.  
 CC -1- ESSENTIAL FOR EXPRESSION OF DCTP (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- PTM: PHOSPHORYLATED BY DCTS (By similarity).  
 CC -1- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AP001516; BAB06470.1;  
 DR PIR: G83993; G83993.  
 DR HSP: P10957; 1RNL.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF00072; response\_reg.1.  
 DR PIRSF: PIRSF006171; RR\_citrat\_malat; 1.  
 DR ProDom: PD000039; Response\_reg.1.  
 DR SMART: SM00448; REC.1.  
 DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.  
 DR Sensory transduction; Transcription regulation; DNA-binding;  
 KW Activator; Phosphorylation; Complete proteome.  
 FT DOMAIN 8 124 RESPONSE REGULATORY.  
 FT MOD\_RES 59 59 PHOSPHORYLATION (BY SIMILARITY).  
 FT DNA\_BIND 183 209 H-T-H MOTIF (POTENTIAL).  
 FT SEQUENCE 230 AA; 26376 MW; 06FD65F46FC57D40 CRC64;  
 Query Match 20.7%; Score 140.5; DB 1; Length 230;  
 Best Local Similarity 29.4%; Pred. No. 6.9e-05;  
 Matches 35; Conservative 32; Mismatches 47; Indels 5; Gaps 3;  
 QY 16 LNVIVDDPLNLIHEKIIKAIG--ISQTANNGEAAVHHRDGGSSFDLIMDKEMPE 73  
 Db 7 IRLVIEDDPVQVFNKQFIITVKGVTCATAGNCEGKLIKE--EQPDVLVDVY 64  
 QY 74 RDGVSTTKKREMEVKSIVGTSLADNEERAFMEAGLNHCLAKPLTKDKIPLINQ 132  
 Db 65 QDGLSFIKQIREQIDVDIIAVTA-ANDTKITLLRYGVMDYLVKPFTEERLKAALTQ 122  
 RESULT 15  
 CHEV\_BACSU STANDARD; PRT: 119 AA.  
 ID CHEV\_BACSU  
 AC P24072; P37583;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chemotaxis protein chev homolog.  
 GN CHEV OR CHEB.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=168 / O11085;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 95 Seconds

(without alignments)  
369.423 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678

Sequence: 1 MATKSMGDIKIKKLNVL.....LAKPLTKDKIPLINQLMDA 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp-archaea:\*

2: sp-bacteria:\*

3: sp-fungi:\*

4: sp-human:\*

5: sp-invertebrate:\*

6: sp-mammal:\*

7: sp-muc:\*

8: sp-organelle:\*

9: sp-phage:\*

10: sp-plant:\*

11: sp-rodent:\*

12: sp-virus:\*

13: sp-vertebrate:\*

14: sp-unclassified:\*

15: sp-ivirius:\*

16: sp-bacteriap:\*

17: sp-archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	136	10	O82445
2	538	79.4	142	10	Q9M8Y4
3	181	26.7	1969	5	O15763 dictyosteli
4	179.5	26.5	820	16	Q8E263
5	178.5	26.3	1197	16	Q8CVU5
6	174	25.7	394	16	Q9A3P0
7	173.5	25.6	949	16	Q8FPF9
8	173	25.5	417	2	Q9RLC7
9	171.5	25.3	933	16	Q8X39
10	171.5	25.3	957	16	Q8ZGR4
11	171	25.2	927	2	Q9ANY0
12	170	25.1	769	16	Q8PQ37
13	167.5	24.7	1364	16	Q8FJN8
14	166.5	24.6	507	2	O85663
15	163.5	24.1	1364	16	Q8F883
16	162.5	24.0	642	16	Q9A3L2

17 162.5 24.0 719 3 Q9P896  
18 162.5 24.0 778 16 Q8ZB69  
19 162 23.9 767 16 Q8PD62  
20 161.5 23.8 148 5 Q9GTU0  
21 161.5 23.8 1709 5 Q9SPH5  
22 161 23.7 716 16 Q8D5I6  
23 160.5 23.7 574 16 Q8D9H9  
24 159.5 23.5 571 16 Q8OAE1  
25 159 23.5 117 17 Q8S8E1  
26 159 23.5 736 16 Q9KRI6  
27 159 23.5 1645 16 Q8YTW3  
28 158.5 23.4 1266 3 Q8NIV2  
29 157 23.2 900 16 Q9AC40  
30 156.5 23.1 469 2 Q8RNK0  
31 156 23.0 1000 17 Q8TPA1  
32 155 22.9 690 16 Q9A501  
33 155 22.9 767 12 Q8QKV7  
34 154.5 22.8 122 17 Q8S8E7  
35 154.5 22.8 572 16 Q9KSI6  
36 154 22.7 708 3 Q9C1Q7  
37 154 22.7 848 2 Q9AEW3  
38 153.5 22.6 713 16 Q9A472  
39 153.5 22.6 939 2 Q8KWS5  
40 153 22.6 1268 2 Q8KQV0  
41 152.5 22.5 120 17 Q48299  
42 152.5 22.5 514 16 Q9ABT2  
43 152.5 22.5 929 16 Q8EBQ2  
44 152.5 22.5 935 2 Q8RMF4  
45 152.5 22.5 1068 16 Q8PMW23

#### ALIGNMENTS

#### RESULT 1

O82445 PRELIMINARY: PRT: 136 AA.  
ID O82445  
AC O82445;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Response regulator protein.  
DE Response regulator protein.  
OS Brassica napus (Rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Whitelaw C.A., Paul W., Jenkins E.S., Taylor V.M., Roberts J.A.;  
RT "A mRNA encoding a response regulator protein from Brassica napus is up-regulated during pod development.";  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF057027; AAC62225.1;  
DR InterPro; IPR001789; Response\_reg.  
DR Pfam; PF00072; Response\_reg.1.  
DR ProDom; PD000039; Response\_reg.1.  
DR SMART; SM00448; REC.1.  
DR PROSITE; PS01110; RESPONSE\_REGULATORY.1.  
KW Phosphorylation; Sensory transduction.  
SQ SEQUENCE 136 AA; 15035 MW; D99B768FAB3CF5 CRC64;

Query Match 100.0%; Score 678; DB 10; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4,7e-52;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATKSMGDIKIKKLNVLIVDDPLNLIIEKIIKAIGGISQTANNNGEEAVIIHRDGS 60  
Db 1 MATKSMGDIKIKKLNVLIVDDPLNLIIEKIIKAIGGISQTANNNGEEAVIIHRDGS 60  
Qy 61 SFDLILMDKEMPERGVSTTKKREMEYKSMIVGVTSLADNEEERRAFMEAGLNHCLAKP 120  
Db 61 SFDLILMDKEMPERGVSTTKKREMEYKSMIVGVTSLADNEEERRAFMEAGLNHCLAKP 120

```

QY 121 LTKDKIPLINQMDA 136
DB 121 LTKDKIPLINQMDA 136
RESULT 2
Q9M8Y4 PRELIMINARY; PRT; 142 AA.
AC Q9M8Y4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative response regulator protein (receiver component).
GN T6K12.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T6K12 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Trouthan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.B.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Braver V., Trouthan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016829; AAF26786.1; -
DR EMBL; AY085638; AAM62859.1; -
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg.1.
DR ProDom; PD000039; Response_reg.1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Phosphorylation; Sensory transduction.
SQ SEQUENCE 142 AA; 13792 MW; 57487135FF2D5238 CRC64;

Query Match 79.4%; Score 538; DB 10; Length 142;
Best Local Similarity 78.9%; Pred. No. 9.8e-40;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 1 MATKSMGDIK-----IKKKL-NVLIVDDPLNLIIHEKIKAIGGISQTANNGEAAVII 54
DB 1 MATKSTGTETKTSIEVKKLNVLIVDDPLNRLRHEWIKTIGGISQTAKNGEAAVII 60
QY 55 HROGGSFLLMDKEMPERDGVSTTKKLREMEVKSMIVGVTSIADNEEERAFMEAGLN 114
DB 61 HRDGEASFLLMDKEMPERDGVSTTKKLREMKVTSIMIVGVTSVADQEEERAFMEAGLN 120
QY 115 HCLAKPLTKDKIPLINQMDA 136
DB 121 HCLEKPLTKAKIPLISHLFDA 142
RESULT 3
O15763 PRELIMINARY; PRT; 1969 AA.
AC O15763;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

```

---

```

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hybrid histidine kinase DHKB.
GN DHKB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAX3;
RX MEDLINE=98248997; PubMed=9576830;
RA Zinda M.J., Singleton C.K.;
RT "The hybrid histidine kinase dhkb regulates spore germination in
RT Dictyostelium discoideum.";
RL Dev. Biol. 196:171-183(1998).
DR EMBL; AF024654; AAB71889.1; -
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_KinA.
DR InterPro; IPR003467; His_Kinase.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HSKA; 1.
DR Pfam; PF00072; response_reg.1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg.1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR TIGRfams; TIGR00229; sensory_box; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 1969 AA; 219025 MW; 857A7952AB1B528 CRC64;

Query Match 26.7%; Score 181; DB 5; Length 1969;
Best Local Similarity 35.8%; Pred. No. 5.1e-07;
Matches 48; Conservative 29; Mismatches 39; Indels 18; Gaps 5;

QY 10 EKIKKLNVLIVDDPLNLIIHEKIKAIGGISQTANNGEAAVIIHRDGGSSFDLILMDK 69
DB 1836 EKIEK---ILVEDNFYVNFVIFSKLLKDSGYIPDVAHNGVEAVECVKKG--AYDILMDK 1890
QY 70 EMERDGVSTTKKLREMEVKSMI-----VGVTSLADNE--EERRAFMEAGLNHCL 117
DB 1891 QMPENDGFEATTAIRELEKSNLIESPPSKKSHVIVALTANSYKDKGKCLSGVMNDFL 1950
QY 118 AKPL-TDKKIPLI 130
DB 1951 QPKITSIDILQMI 1964
RESULT 4
Q8EZ63 PRELIMINARY; PRT; 820 AA.
AC Q8EZ63;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN LA3996.
OS Leptosira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptosiraceae; Leptosira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011554; AANS1193.1; -
KW Complete proteome.
SQ SEQUENCE 820 AA; 94325 MW; 3DBE99966794E67D CRC64;

```

```

Query Match          26.5%; Score 179.5; DB 15; Length 820;
Best Local Similarity 31.6%; Pred. No. 2.4e-07;
Matches 43; Conservative 29; Mismatches 49; Indels 15; Gaps 5;

QY 8 DIEK-----IKKLNVLIVDDPLNLIIEKIIKAIGGISQTANNGEAVIIHRDGGSGF 62
DB 679 DLEKSFQDVVAKSTRILVAEDNETNCLLIERALKKLGVDPTVVHNGRE--VIERMQLEAF 736
QY 63 DLILMDKEMPERGVSTTKKIREMEVKS---MIVGVTSIADNEEERRAFMEAGLNHCLAK 119
DB 737 DIILMDHMPEDVIGATKWRKSNQNSFEPIIIALTADA-ISSKERYIKSGMNDCLTK 795
QY 120 PLTKDKIPLINLMD 135
DB 796 PLD-----LPLKSTLD 807

RESULT 5
Q8CVU5 PRELIMINARY; PRT: 1197 AA.
AC Q8CVU5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Sensor protein evgs precursor (EC 2.7.3.-).
GN EVGS OR C2906.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.J., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016764; AA81356.1; -
KW Transferase; Complete proteome.
SQ SEQUENCE 1197 AA; 134847 MW; 5088214E7F834F06 CRC64;

Query Match          26.38; Score 178.5; DB 16; Length 1197;
Best Local Similarity 34.4%; Pred. No. 4.6e-07;
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;

QY 12 IKKLNVLIVDDPLNLIIEKIIKAIGGISQTANNGEAVIIHRDGGSGFLLMDKEM 71
DB 955 LPKXLSILIAADHPNTNLLKRLQLNLGVDVDEATGVQA--LHKYSMQHYDLLITDVNM 1012
QY 72 PERGVSTTKKIREMEVKSIVGVTSIADNEEERRAFMEAGLNHCLAKPLTKDKIPLIN 131
DB 1013 PNMDFELTRKLRQNSLSPIWGLTANA-QANEREXGLNGMNLCLFKPLTLDVLKTHLS 1071
QY 132 QL 133
DB 1072 QL 1073

RESULT 6
Q9A3P0 PRELIMINARY; PRT: 394 AA.
AC Q9A3P0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Sensor histidine kinase/response regulator.
GN CC3162.
OS Caulobacter crescentus.

```

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit K., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
DR EMBL; AE005980; AA25124.1; -
DR HSP; P06657; 2CHF.
DR TIGR; CC3162; -.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c.1.
DR Pfam; PF00072; response_reg.1.
DR ProDom; PD000039; Response_reg.1.
DR SMART; SM00387; HATPase_c.1.
DR SMART; SM00448; REC.1.
DR PROSITE; PS01019; HIS_KIN.1.
DR PROSITE; PS01110; RESPONSE_REGULATORY.1.
KW Kinase; Phosphorylation; Sensory transduction; Complete proteome.
SQ SEQUENCE 394 AA; 41592 MW; 6D9AF0B278B70E70 CRC64;

Query Match          25.7%; Score 174; DB 16; Length 394;
Best Local Similarity 31.7%; Pred. No. 3.1e-07;
Matches 39; Conservative 31; Mismatches 49; Indels 4; Gaps 2;

QY 14 KKLNVLIIVDDPLNLIIEKIIKAIGGISQTANNGEAVIIHRDGGSGFLLMDKEMPE 73
DB 262 RSAHLIVDDNATNRVVAEALCDMFCTSEQAVDGVAEVEMARSG--RFDLILMDIKMPR 319
QY 74 RDGVSTTKKIREMEVKSIVGVTSIADNEE--ERRAFMEAGLNHCLAKPLTKDKIPLIN 131
DB 320 MDGVAATRAIRELSGRSSAAPIVALTANADPADVHTYLAAGMDVYVEKPIKPERLALV 379
QY 132 QLM 134
DB 380 SLL 382

RESULT 7
Q8FFP9 PRELIMINARY; PRT: 949 AA.
AC Q8FFP9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sensor protein rscs (EC 2.7.3.-).
GN RSCS OR C2761.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.J., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016764; AA81356.1; -
KW Transferase; Complete proteome.
SQ SEQUENCE 1197 AA; 134847 MW; 5088214E7F834F06 CRC64;

```

RT	of uropathogenic Escherichia coli.,"	Query Match	25.6%	Score 173.5	DB 16	Length 949
RL	PROC. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).	Best Local Similarity	33.6%	Pred. No. 9.7e-07		
DR	EMBL; AE016763; AAN81215.1; -	Matches	37	Conservative	35	Mismatches 33; Indels 5; Gaps 3;
SR	Transferrase; Complete proteome.					
SQ	SEQUENCE 949 AA; 106590 MW; DF8CA47F9EEB4088 CRC64;					
QY	18 VLIIVDDPLNIIIEKIIKAIGISQTANGGEAV-IHRDGGSSFDLILMDKEMPERDG 76					
DB	827 ILVVDHPINRRLADGLSGYQCKTANGDVALNLYNK---NHIDIVLSYNNPNMDG 883					
QY	77 VSTTKRLREMEVKSVMIGVTSLADNSEEERRAFREAGLNHCLAKPTKDKI 126					
DB	884 YELTORIQGLGLTPVIGVTANALAEKQRC-LESGMDSCLSPVTLDVI 932					
RESULT 8						
Q9RLC7	PRELIMINARY; PRT; 417 AA.					
ID	Q9RLC7					
AC	Q9RLC7					
DT	01-MAY-2000 (TrEMBLrel. 13, Created)					
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)					
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)					
DE	Putative histidine kinase (Fragment).					
GN	GACS.					
OS	Pseudomonas stutzeri (Pseudomonas perfectomarina).					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;					
OC	Pseudomonadaceae; Pseudomonas.					
NCBI_TAXID=316;						
[1]	SEQUENCE FROM N.A.					
RN	STRAIN-JM300;					
RC	Graupner S.; Wackernagel W.;					
RA	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; AJ249741; CAB56474.1; -					
DR	InterPro; IPR002570; Hpt.					
DR	InterPro; IPR001789; Response_reg.					
DR	Pfam; PF00072; response_reg.1.					
DR	ProDom; PD000039; Response_reg.1.					
DR	SMART; SM00073; HPT.1.					
DR	SMART; SM00448; REC.1.					
DR	PROSITE; PS50110; RESPONSE_REGULATORY.1.					
KW	Kinase; Phosphorylation; Sensory transduction.					
FT	NON_TER 1					
SQ	SEQUENCE 417 AA; 45749 MW; D196B4FFDDA23B32 CRC64;					
Query Match	25.5%	Score 173;	DB 2;	Length 417;		
Best Local Similarity	32.8%	Pred. No. 4e-07;				
Matches	39;	Conservative	35;	Mismatches	37;	Indels 8; Gaps 3
QY	17 NVLIIVDDPLNIIIEKIIKAIGISQTANGGEAV-IHRDGGSSFDLILMDKEMPERDG 76					
DB	169 SVLCVDDNPNALMLLETLLTDMGGEVAVSSGQALVYVQ--QSFDMVFMVQNPFGMDG 226					
QY	77 VSTTKRLREMEVKSVMIGVTSLADNSEEERRAFREAGLNHCLAKPTKDKIPLI 130					
DB	227 RQTTEAIRNWELESQGPPLPIVALTAHA-LSNERRSLQSLDDYLTKPISERQLAQV 284					
RESULT 9						
Q8XE39	PRELIMINARY; PRT; 933 AA.					
ID	Q8XE39					
AC	Q8XE39;					
DT	01-MAR-2002 (TrEMBLrel. 20, Created)					
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)					
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)					
DE	Sensor for ctr capsule biosynthesis, probable histidine kinase acting on RcsB.					
GN	RCSB OR 23477 OR ECS3107.					
OS	Escherichia coli O157:H7.					

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Versinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;  
RT "Genome sequence of *Versinia pestis*, the causative agent of plague";  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIM5 / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Retherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of *Versinia pestis* KIM";  
RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL: AJ414147; CAC90055.1; -;  
DR EMBL: AE013898; AM86522.1; -;  
DR HSSP: P06143; 1D42.  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003661; His\_kinase.  
DR InterPro: IPR003467; His\_kinase.  
DR InterPro: IPR001789; Response\_reg.  
DR Pfam: PF02518; HSKA; 1.  
DR Pfam: PF00072; response\_reg; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
DR PROSITE: PS50110; RESPONSE\_REGULATORY; 1.  
DR Kinase; Transferase; Complete proteome.  
KW Kinase; Transferase.  
SQ SEQUENCE 957 AA; 108591 MW; 0447A11F59100011 CRC64;  
  
Query Match 25.3%; Score 171.5; DB 16; Length 957;  
Best Local Similarity 35.1%; Pred. No. 1.5e-06;  
Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;  
  
QY 16 LNVLIYDDPLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIIMDKEMPERD 75  
Db 835 LQILVYDDHPINRLADQLTGLYRVITANDGLDALVNT--NTVDMVLTVDNPNND 892  
QY 76 GVSTTKKLEMEYKSMIVGTVSLADNEERRAFMEAGLNHCLAKPLTKDKI 126  
Db 893 GYRLTERLQNLHNPFIIIGVTANALAEKGQRC-IEAGMDNCLSKPVTLDTL 942  
  
RESULT 11  
Q9ANYO PRELIMINARY; PRT; 927 AA.  
AC Q9ANYO;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE RSCS.  
GN RSCS.  
OS *Vibrio fischeri*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=668;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=21142508; PubMed=11208780;  
RA Visick K.L., Skoufos L.M.;  
RT "A two-component sensor required for normal symbiotic colonization of  
RT *Euprymna scolopes* by *Vibrio fischeri*.";  
RL J. Bacteriol. 183:835-842(2001).  
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE  
CC KINASES  
DR EMBL: AF319618; AAG60694.1; -;  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003661; His\_kinase.  
DR InterPro: IPR005467; His\_kinase.  
DR InterPro: IPR002570; Hpt.  
DR InterPro: IPR001610; PAC.  
DR InterPro: IPR000700; PAS-assoc\_C.  
DR InterPro: IPR000014; PAS\_domain.  
DR InterPro: IPR001789; Response\_reg.  
DR Pfam: PF02518; HSKA; 1.  
DR Pfam: PF01627; Hpt; 1.  
DR Pfam: PF00785; PAC; 1.  
DR Pfam: PF00989; PAS; 1.  
DR Pfam: PF00072; response\_reg; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR PROSITE: PS50112; HSKA; 1.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; HSKA; 1.  
DR SMART: SM00073; HPT; 1.  
DR SMART: SM00086; PAC; 1.  
DR SMART: SM00091; PAS; 1.  
DR SMART: SM00448; REC; 1.  
DR TIGRFAMs: TIGR00229; sensory\_box; 1.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
DR PROSITE: PS50113; PAC; 1.  
DR PROSITE: PS50112; PAS; 1.  
DR PROSITE: PS50110; RESPONSE\_REGULATORY; 1.  
DR Kinase; Phosphorylation; Sensory transduction; Transferase.  
KW Kinase; Phosphorylation.  
SQ SEQUENCE 927 AA; 105590 MW; 9EE059CB2E8BF8B CRC64;  
  
Query Match 25.2%; Score 171; DB 2; Length 927;  
Best Local Similarity 33.9%; Pred. No. 1.6e-06;  
Matches 43; Conservative 28; Mismatches 42; Indels 14; Gaps 4;  
  
QY 16 LNVLIYDDPLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIIMDKEMPERD 75  
Db 659 LSILTAEDNKINALVAQMFQRLGHKAIENGK--VAIDKLETHFDLIIMDNHMPVD 716  
QY 76 GYSTTKKLEMEYKSMIVGTVSLA-----DNEERRAFMEAGLNHCLAKPLTKDKIPL 129  
Db 717 GILATKIIREKLIKISTVIFAYTANAFQAHDN-----FLKAGANYVLTPLQENDEFIGA 770  
QY 130 INQLMDA 136  
Db 771 IKQYQDA 777  
  
RESULT 12  
Q8PQ37 PRELIMINARY; PRT; 769 AA.  
AC Q8PQ37;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Two-component system sensor protein.  
GN XAC0494.  
OS *Xanthomonas axonopodis* (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; *Xanthomonas*.  
OX NCBI\_TaxID=92829;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=306 / ATCC 13902 / XV 101;  
RL MEDLINE=22022145; PubMed=12024217;  
RX

RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA	Martins E.C., Magdano J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA	Moraes L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA	Spinola L.A.F., Takita M.A., Tamara R.E., Teixeira E.C., Tezza R.I.D.,
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA	Subetal J.C., Kitajima J.P.;
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
RT	host specificities";
RL	Nature 417:459-463(2002).
DR	EMBL; AE011887; AAM37343.1; -
DR	InterPro; IPR003594; ATPbind_ATPase.
DR	InterPro; IPR004358; bact_sens_pr_C.
DR	InterPro; IPR006189; CHASE.
DR	InterPro; IPR003661; his_kinA.
DR	InterPro; IPR005467; his_kinase.
DR	InterPro; IPR002570; hpt.
DR	InterPro; IPR001610; PAC.
DR	InterPro; IPR000700; PAS-assoc_C.
DR	InterPro; IPR000014; PAS_domain.
DR	InterPro; IPR001789; Response_reg.
DR	Pfam; PF03924; CHASE; 1.
DR	Pfam; PF02518; HATPase_c; 1.
DR	Pfam; PF00512; HisKA; 1.
DR	Pfam; PF01627; Hpt; 1.
DR	Pfam; PF00785; PAC; 2.
DR	Pfam; PF00985; PAS; 2.
DR	Pfam; PF00072; response_reg; 2.
DR	PRINTS; P000344; BCTRSENSOR.
DR	PRODOM; PD000039; Response_reg; 2.
DR	TIGRFAMS; TIGR00229; sensory_box; 2.
DR	PROSITE; PS50839; CHASE; 1.
DR	PROSITE; PS50109; HIS_KIN; 1.
DR	PROSITE; PS50113; PAC; 2.
DR	PROSITE; PS50112; PAS; 2.
DR	PROSITE; PS50110; RESPONSE_REGULATORY; 2.
DR	Complete proteome.
KW	SEQUENCE 1364 AA; 150282 MW; F7CCA3856B7E7165 CRC64;
SQ	
Query Match	24.7%; Score 167.5; DB 16; Length 1364;
Best Local Similarity	29.2%; Pred. No. 5e-06;
Matches 35; Conservative	37; Mismatches 45; Indels 3; Gaps 2;
QY	17 NVLIVDDPLNLIHKKIIKAGGISQTANNKEAVIHRDGGSSFDLLMDKENPEDG 76 :::     :  ::    :  :  ::    :  :  ::    :  :  ::    :  :
Db	1033 HLLVYDSDINCVAQRILSEGAVTVAHDGEQAVSTLKRAPNFLHLVMDVQVPVVG 1092
QY	77 VSTTKKLREMEVKSM--IVGVTSIADNEERRAFMEAGLNHCIAKPLTKDKIIPNLQM 134  :::    :
Db	1093 YEATFRLRQPALASLPVIALTAGARFPQOKEA-LEAGMNGFIAPFNVEELVTAIRHFL 1151
RESULT 14	
O85663	PRELIMINARY; PRT; 507 AA.
ID O85663	
AC O85663;	
DT 01-NOV-1998 (TReMBLrel. 08, Created)	
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)	
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
DE RCSC (Fragment).	
GN RCSC.	
OS Proteus mirabilis.	
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC Enterobacteriaceae; Proteus.	
OX NCBI_TaxID=584;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-BB2000;	
RX MEDLINE=99047557; PubMed=9829920;	
RA Belas R., Schneider R., Melch M.;	
RT "Characterization of Proteus mirabilis precocious swarming mutants:	
RT identification of tsba, encoding a regulator of swarming behavior.";	

```

RL J. Bacteriol. 180:6126-6139(1998).
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AF071215; AAC82662.1; -.
DR HSSP: P06657; 2CHF.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pt_C.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; HSKA; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00386; HSKA; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
DR Kinase: Phosphorylation; Sensory transduction; Transferase.
KW NON_TER
FT 1
SQ SEQUENCE 507 AA; 57326 MW; F5DA2EAD9C35DEF5 CRC64;

Query Match 24.6%; Score 166.5; DB 2; Length 507;
Best Local Similarity 36.4%; Pred. No. 1.9e-06;
Matches 39; Conservative 26; Mismatches 39; Indels 3; Gaps 2;

QY 16 LNVLIYDDPLNLIHKKIKAIGISQTANNGEAVIHRDGGSSFDLILMDKEMPERD 75
DB 397 LTVLIYDDHPFLRLTDLQKIGENTAFEDGCDALAFMQE--NHVDIILTDVNMNPN 454

QY 76 GVSTTKKLEMEVKSIMVGTSLADNEBERRAFMEAGLNHCLAKPLT 122
DB 455 GYQIATVRELSSTIPIIGVTANATAEAKQRC-IDAGMNDVCVSPVS 500

RESULT 15
Q8P883 PRELIMINARY; PRT; 1364 AA.
AC Q8P883;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component system sensor protein.
GN XCC2360.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; Pubmed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.; of two Xanthomonas pathogens with differing
RT host specificities.
RL Nature 417:459-463(2002).
EMBL: AE012344; AAM41638.1; -.

```

```

DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pt_C.
DR InterPro: IPR006189; CHASE.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR002570; Hpt.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-associ_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF03924; CHASE; 1.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; HSKA; 1.
DR Pfam: PF01627; Hdc; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 2.
DR PRINTS: PR00344; BCTRLSENSOR.
DR TIGRFAM: TIGR00229; sensory_box; 2.
DR PROSITE: PS50839; CHASE; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50113; PAC; 2.
DR PROSITE: PS50112; PAS; 2.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 2.
KW Complete proteome.
SQ SEQUENCE 1364 AA; 150167 MW; 48F1C5FBD9710316 CRC64;

Query Match 24.1%; Score 163.5; DB 16; Length 1364;
Best Local Similarity 30.3%; Pred. No. 1.1e-05;
Matches 36; Conservative 36; Mismatches 44; Indels 3; Gaps 3;

QY 18 VLIYDDPLNLIHKKIKAIGISQTANNGEAVIHRDGGSSFDLILMDKEMPERDGV 77
DB 1034 LLIYDDSEINCEVAQRILEGEGAMVTVAHDGQAVNTLKRAPDLFHLVIMDVQMPVVDGY 1093

QY 78 STTKKLEME-EVKSM-IVGVTSIADNEBERRAFMEAGLNHCLAKPLTDXDKIIPLINQLM 134
DB 1094 EATRRLRQPSLASLPVIALTAGAFRPOQEA-LEAGMNGFIKPFNVVELVTAIRHFL 1151

Search completed: August 14, 2003, 16:49:03
Job time : 98 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2003, 19:06:41 ; Search time 41 Seconds  
(without alignments)  
536.508 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678

Sequence: 1 MATKSMGDIKIKKLNVL.....LAKPLTKDKIPLINQLMDA 136

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	678	100.0	136	AAV42637	Brassica napus
2	678	100.0	136	AAV42637	Brassica napus
3	538	79.4	142	AA10549	Arabidopsis thaliana
4	538	79.4	142	AA10549	Arabidopsis thaliana
5	515	76.0	142	AAV42645	A. thaliana D22AT3
6	472	69.6	116	AAV42652	Brassica napus
7	437	64.5	104	AA10550	Arabidopsis thaliana
8	437	64.5	104	AA10550	Arabidopsis thaliana
9	301	44.4	70	AA10589	Arabidopsis thaliana

10	280	41.3	66	21	AA10590	Arabidopsis thaliana
11	213	31.4	71	21	AA10588	Arabidopsis thaliana
12	202	29.8	261	21	AA10588	Arabidopsis thaliana
13	192.5	28.4	67	20	AA10588	Arabidopsis thaliana
14	171.5	25.3	118	23	AA10588	Arabidopsis thaliana
15	156.5	23.1	471	24	AA10588	Arabidopsis thaliana
16	156.5	23.1	496	24	AA10588	Arabidopsis thaliana
17	156.5	23.1	666	24	AA10588	Arabidopsis thaliana
18	156.5	23.1	709	24	AA10588	Arabidopsis thaliana
19	156.5	23.1	719	24	AA10588	Arabidopsis thaliana
20	156.5	23.1	802	24	AA10588	Arabidopsis thaliana
21	156.5	23.1	870	24	AA10588	Arabidopsis thaliana
22	156.5	23.1	887	24	AA10588	Arabidopsis thaliana
23	156.5	23.1	900	24	AA10588	Arabidopsis thaliana
24	156.5	23.1	906	24	AA10588	Arabidopsis thaliana
25	156.5	23.1	931	24	AA10588	Arabidopsis thaliana
26	156.5	23.1	922	24	AA10588	Arabidopsis thaliana
27	156.5	23.1	922	24	AA10588	Arabidopsis thaliana
28	149	22.0	1373	23	AA10588	Arabidopsis thaliana
29	148.5	21.9	922	22	AA10588	Arabidopsis thaliana
30	148.5	21.9	922	22	AA10588	Arabidopsis thaliana
31	148.5	21.9	950	23	AA10588	Arabidopsis thaliana
32	146	21.5	1018	21	AA10588	Arabidopsis thaliana
33	141	20.8	1146	20	AA10588	Arabidopsis thaliana
34	139.5	20.6	1281	20	AA10588	Arabidopsis thaliana
35	139.5	20.6	1298	19	AA10588	Arabidopsis thaliana
36	139.5	20.6	1298	20	AA10588	Arabidopsis thaliana
37	139.5	20.6	1298	20	AA10588	Arabidopsis thaliana
38	137	20.2	125	23	AA10588	Arabidopsis thaliana
39	134.5	19.8	1122	23	AA10588	Arabidopsis thaliana
40	134	19.8	120	22	AA10588	Arabidopsis thaliana
41	133.5	19.7	1023	22	AA10588	Arabidopsis thaliana
42	131	19.3	136	23	AA10588	Arabidopsis thaliana
43	131	19.3	674	21	AA10588	Arabidopsis thaliana
44	130	19.2	1081	20	AA10588	Arabidopsis thaliana
45	130	19.2	1081	20	AA10588	Arabidopsis thaliana

## ALIGNMENTS

RESULT 1  
AAV42637

ID AAV42637 standard; Protein; 136 AA.

AC AAV42637;

XX 10-JAN-2000 (first entry)

XX Brassica napus response regulator protein D22.

XX Signal transduction protein; dehiscence; male sterile plant; D22 gene;

XX Shatter resistance; oilseed rape; response regulator protein.

XX Brassica napus.

XX WO9949046-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-GB00905.

XX 20-MAR-1998; 98GB-0006113.

XX (BIOG-) BIOEMMA UK LTD.

XX Wyatt P, Roberts JA, Whitelaw C;

XX WPI; 1999-580449/49.

XX N-PSDB; AAZ22974.

XX A nucleic acid encoding a signal transduction protein involved in plant dehiscence, useful for producing shatter resistant male sterile plants

PT  
XX  
PS  
XX  
XX  
Claim 4; Fig 1; 7lpp; English.

CC The invention provides a nucleic acid encoding a signal transduction protein involved in the process of dehiscence. The nucleic acids and CC proteins are useful for regulating or controlling dehiscence of a pod or CC an anther in a plant, useful in the production of male sterile plants. CC The methods, etc. may be used in the production of shatter resistance or CC shatter-delayed plants such as oilseed rape (*Brassica napus*). The present CC sequence represents a B. napus response regulator protein DZ2.

XX  
XX  
Sequence 136 AA;

Query Match 100.0%; Score 678; DB 20; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1e-67;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKSMGDIKKKLVIVDDPLNLIHKKIKAIGGISQTANNGEAAVTHRDGGS 60  
DB 1 MATKSMGDIKKKLVIVDDPLNLIHKKIKAIGGISQTANNGEAAVTHRDGGS 60

QY 61 SFDLILMDKEMPERDGVSTTKKLRMEVKSMIVGVTSADNEEERRAFMEAGLNHCLAKP 120  
DB 61 SFDLILMDKEMPERDGVSTTKKLRMEVKSMIVGVTSADNEEERRAFMEAGLNHCLAKP 120

QY 121 LTKDKIIPLINQMDA 136  
DB 121 LTKDKIIPLINQMDA 136

RESULT 2  
AA42653  
ID AA42653 standard; Protein: 136 AA.  
XX  
AC AA42653;  
XX  
DT 10-JAN-2000 (first entry)  
XX  
DE Brassica response regulator protein DZ2 putative peptide sequence.  
XX  
KW Signal transduction protein; dehiscence; male sterile plant; DZ2 gene;  
XX shatter resistance; oilseed rape; response regulator protein.  
XX  
OS Brassica napus.  
XX WO949046-A1.  
XX  
XX  
XX 30-SEP-1999.  
XX  
XX 22-MAR-1999; 99WO-GB00905.  
XX  
XX 20-MAR-1998; 98GB-0006113.  
XX  
XX (BIOG-) BIOGEMMA UK LTD.  
XX  
XX Wyatt P, Roberts JA, Whitelaw C;  
XX  
XX WPI; 1999-580449/49.  
XX  
XX N-PSDB; AA222977.  
XX  
XX A nucleic acid encoding a signal transduction protein involved in plant  
PT dehiscence, useful for producing shatter resistant male sterile plants  
PT  
PT  
XX  
XX Example 2; Fig 6; 7lpp; English.

CC The invention provides a nucleic acid encoding a signal transduction CC protein involved in the process of dehiscence. The nucleic acids and CC proteins are useful for regulating or controlling dehiscence of a pod or CC an anther in a plant, useful in the production of male sterile plants. CC The methods, etc. may be used in the production of shatter resistance or CC shatter-delayed plants such as oilseed rape (*Brassica napus*). The present

CC sequence represents the B. napus DZ2 putative peptide sequence.

XX  
XX  
XX  
Sequence 136 AA;

Query Match 100.0%; Score 678; DB 20; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1e-67;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKSMGDIKKKLVIVDDPLNLIHKKIKAIGGISQTANNGEAAVTHRDGGS 60  
DB 1 MATKSMGDIKKKLVIVDDPLNLIHKKIKAIGGISQTANNGEAAVTHRDGGS 60

QY 61 SFDLILMDKEMPERDGVSTTKKLRMEVKSMIVGVTSADNEEERRAFMEAGLNHCLAKP 120  
DB 61 SFDLILMDKEMPERDGVSTTKKLRMEVKSMIVGVTSADNEEERRAFMEAGLNHCLAKP 120

QY 121 LTKDKIIPLINQMDA 136  
DB 121 LTKDKIIPLINQMDA 136

RESULT 3  
AAG10549  
ID AAG10549 standard; Protein: 142 AA.  
XX  
AC AAG10549;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8916.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131448.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.

PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138647.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-0139493.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149829.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149802.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	18-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	21-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	22-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140931.	PR	15-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0142154.	PR	22-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-0155486.
PR	06-JUL-1999;	99US-0142390.	PR	24-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142820.	PR	29-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	06-OCT-1999;	99US-0158232.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144085.	PR	08-OCT-1999;	99US-0158369.
PR	16-JUL-1999;	99US-0144086.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144325.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159330.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144332.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	21-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160815.
PR	22-JUL-1999;	99US-0145152.	PR	21-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160989.
PR	23-JUL-1999;	99US-0145224.	PR	25-OCT-1999;	99US-0161404.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145918.	PR	26-OCT-1999;	99US-0161359.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161360.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.			

PR	28-OCT-1999;	99US-0161920.	PR	20-MAY-1999;	99US-0135124
PR	28-OCT-1999;	99US-0161920.	PR	21-MAY-1999;	99US-0135353
PR	28-OCT-1999;	99US-0161992.	PR	24-MAY-1999;	99US-0135629
PR	28-OCT-1999;	99US-0161993.	PR	25-MAY-1999;	99US-0136021
PR	29-OCT-1999;	99US-0162142.	PR	27-MAY-1999;	99US-0136392
Query Match 79.4%; Score 538; DB 21; Length 142;					
Best Local Similarity 78.9%; Pred. No. 4.7e-52;					
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;					
QY	1 MATKSMGDIEK----	IKKKL-NVLIYDDPLNLIHKKI	KAIGISQTANNGBEAVII	54	
DB	1 MATKSTGTEKTSIEVKKKLN	VIYDDPLNRLHLEMIITGIGISQTAKNGEAVIL	60		
QY	55 HRDGSFLLIMKEMPERDGVST	TKKLEMEVKSMIVGVTSLADNEEERRAFMEAGLN	114		
DB	61 HRDGEASFLLIMKEMPERDGVST	TKKLEMKVTSMIVGVTSVADQOEERKAFMEAGLN	120		
QY	115 HCLAKPLTKRIIPLINQMDA	136			
DB	121 HCLEKPLTKAKIFPLISHLEDA	142			
RESULT 4					
AAG44439					
ID	AAG44439 standard; Protein; 142 AA.				
XX	AC				
AC	AAG44439;				
XX	DT				
XX	18-OCT-2000 (first entry)				
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 55668.				
DE	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KX	termination sequence.				
KW	Arabidopsis thaliana.				
XX	Arabidopsis thaliana.				
OS	EP1033405-A2.				
XX	06-SEP-2000.				
PN	25-FEB-2000; 2000EP-0301439.				
PD	25-FEB-1999;				
PF	99US-0121825;				
PF	99US-0123160;				
PR	09-MAR-1999;				
PR	99US-0125548;				
PR	99US-0125786;				
PR	23-MAR-1999;				
PR	99US-0126264;				
PR	29-MAR-1999;				
PR	99US-0126785;				
PR	01-APR-1999;				
PR	99US-0127462;				
PR	06-APR-1999;				
PR	99US-0128234;				
PR	08-APR-1999;				
PR	99US-0128714;				
PR	16-APR-1999;				
PR	99US-0129845;				
PR	19-APR-1999;				
PR	99US-0130077;				
PR	21-APR-1999;				
PR	99US-0130449;				
PR	23-APR-1999;				
PR	99US-0130510;				
PR	28-APR-1999;				
PR	99US-0130891;				
PR	30-APR-1999;				
PR	99US-0131449;				
PR	04-MAY-19				

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152863.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158028.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 79.4%; Score 538; DB 21; Length 142;
Best Local Similarity 78.9%; Pred. No. 4.7e-52;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 1 MATKSMGDIKK-----IKKKL-NVLIIVDDDDPLNLIHKKIKAIGGISTQANNGEAVII 54
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MATKSTGGTEKTSIEVKKKLNLVLIIVDDDDPLNRLHEMIKIITGGISTQAKNGEEAVIL 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 55 HRDGGSSFDLILMDKEMPERGVSTTKKLREMEYKSMIVGVTSIADNEERRAFMEAGLN 114
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 HRDGEASFLLILMDKEMPERGVSTTKKLREMKVTSIMIVGVTSVADQEEERKAFMEAGLN 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 115 HCLAKPLTKDKIPLINOLMDA 136
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 HCLEKPLTKAKIFPLISHLEDA 142
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 5
AAV42645
ID AAV42645 standard; Protein; 142 AA.
XX AAV42645;
AC AAV42645;
XX 10-JAN-2000 (first entry)
DT 10-JAN-2000 (first entry)
XX A. thaliana DZ2AT3 putative peptide sequence.
DE A. thaliana DZ2AT3 putative peptide sequence.
XX Signal transduction protein; dehiscence; male sterile plant;
KW shatter resistance; oilseed rape; DZ2AT3 gene.
KW Arabidopsis thaliana.
XX
OS
XX Key Location/Qualifiers
FH Misc-difference 57 /note= "unknown"
FT Misc-difference 88 /note= "unknown"
FT Misc-difference 88 /note= "unknown"
FT
XX WO9949046-A1.
PN
XX 30-SEP-1999.
PD
XX 22-MAR-1999; 99WO-GB00905.
PF
XX 20-MAR-1998; 98GB-0006113.
PR
XX (BIOG-) BIOGEMMA UK LTD.
PA
XX Wyatt P, Roberts JA, Whitelaw C;
PI WPI: 1999-580449/49.
DR N-PSDB; AAZ22978.
XX
XX A nucleic acid encoding a signal transduction protein involved in plant
PT dehiscence, useful for producing shatter resistant male sterile plants
PT -
XX
XX Example 3; Fig 9; 71pp; English.
PS
XX The invention provides a nucleic acid encoding a signal transduction
CC protein involved in the process of dehiscence. The nucleic acids and
CC proteins are useful for regulating or controlling dehiscence of a pod or
CC an anther in a plant, useful in the production of male sterile plants.
CC The methods, etc. may be used in production of shatter resistance or
CC shatter-delayed plants such as oilseed rape (Brassica napus). The present
CC sequence represents a DZ2 A. thaliana homologue DZ2AT3 putative peptide
CC sequence.
```

```

XX SQ Sequence 142 AA;
Query Match 76.0%; Score 515; DB 20; Length 142;
Best Local Similarity 76.1%; Pred. NO. 1.7e-49;
Matches 108; Conservative 8; Mismatches 20; Indels 6; Gaps 2;

QY 1 MATKSMGDIK-----TKKLL-NVLIVDDPLNLIHEKIIKAIGGISQTANNGEEAVII 54
DB 1 MATKSTGTEKTSIEVKKLLINVLIVDDPLNRLHEMIITIGISQTAKNGEEVIL 60
QY 55 HRDGGSSFDLILMDKEMPERDGVSTTKKREMEVKSMIVGVTSLADNEEERRAFMEAGLN 114
DB 61 HRDGEASFDLILMDKEMPERDGVSTTKKREMKGTSMIVGVTSVADQBEERKAFMEAGLN 120
QY 115 HCLAKPLTKDKIPLINOLMDA 136
DB 121 HCLEKPLTKAKIPPLISHLEDA 142

RESULT 6
AAY42652
ID AAY42652 standard; Protein; 116 AA.
AC AAY42652;
XX
XX
DT 10-JAN-2000 (first entry)
DE Brassica response regulator protein D22B putative peptide sequence.
KW Signal transduction protein; dehiscence; male sterile plant; D22B gene;
KW shatter resistance; oilseed rape; response regulator protein.
XX Brassica napus.
XX
XX Key Location/Qualifiers
FH Misc-difference 2 /note= "unknown"
FT
XX WO9949046-A1.
XX
XX 30-SEP-1999.
XX
XX 22-MAR-1999; 99WO-GB00905.
XX
XX 20-MAR-1998; 98GB-0006113.
XX
XX (BIOG-) BIOGEMMA UK LTD.
XX
XX Wyatt P, Roberts JA, Whitelaw C;
XX
XX WPI: 1999-580449/49.
XX N-PSDB; AAZ22976.
XX
XX A nucleic acid encoding a signal transduction protein involved in plant
XX dehiscence, useful for producing shatter resistant male sterile plants
XX
XX Example 2; Fig 6; 71pp; English.
XX
XX The invention provides a nucleic acid encoding a signal transduction
XX protein involved in the process of dehiscence. The nucleic acids and
XX proteins are useful for regulating or controlling dehiscence of a pod or
XX an anther in a plant, useful in the production of male sterile plants.
XX The methods, etc. may be used in production of shatter resistance or
XX shatter-delayed plants such as oilseed rape (Brassica napus). The present
XX sequence represents the B. napus D22B putative peptide sequence.
XX
XX Sequence 116 AA;
Query Match 59.6%; Score 472; DB 20; Length 116;
Best Local Similarity 80.2%; Pred. NO. 8.4e-45;
Matches 93; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 21 VDDPLNLIHEKIIKAIGGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERDGVSTT 80
DB 1 VDDDFVIRKLHEIIIIKISIGISQTAKNGEEAVNHRDGNASFDLILMDKEMPERDGLSAT 60
QY 81 KKLREMEVKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKDKIPLINOLMDA 136
DB 61 KKLREMKVTSIIIGVTTADNEEERKAFMEAGLNHCLAKPLSKAKIPLINNLMDA 116

RESULT 7
AAG10350
ID AAG10350 standard; Protein; 104 AA.
XX
XX AAG10350;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8917.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 08-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.

```

PR	10-JUN-1999;	99US-01388647.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-01391119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149829.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150366.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139889.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155119.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144008.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159283.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144336.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144684.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-014			

Db 2 IIKTIGISOTAKNGEAVILHRDGEAFDILIMDKEMPERDGVSTTKKLEMKVTSIV 61  
QY 94 GVTSLADNEERAFMEAGLNHCLAKPLTKDKIILINOLMDA 136  
Db 62 GVTSDQDEERAFMEAGLNHCLAKPLTKAKIFPLSHLEDA 104

RESULT 8  
ID AAG44440 standard; Protein; 104 AA.  
XX AAG44440;  
AC AAG44440;  
XX  
DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55669.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128024.  
PR 08-APR-1999; 99US-0128214.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 99US-0139452.

16-JUN-1999; 99US-0139453.  
17-JUN-1999; 99US-0139452.  
18-JUN-1999; 99US-0139454.  
18-JUN-1999; 99US-0139455.  
18-JUN-1999; 99US-0139456.  
18-JUN-1999; 99US-0139457.  
18-JUN-1999; 99US-0139458.  
18-JUN-1999; 99US-0139459.  
18-JUN-1999; 99US-0139460.  
18-JUN-1999; 99US-0139461.  
18-JUN-1999; 99US-0139462.  
18-JUN-1999; 99US-0139463.  
18-JUN-1999; 99US-0139750.  
18-JUN-1999; 99US-0139763.  
21-JUN-1999; 99US-0139817.  
22-JUN-1999; 99US-0139899.  
23-JUN-1999; 99US-0140353.  
23-JUN-1999; 99US-0140354.  
24-JUN-1999; 99US-0140695.  
28-JUN-1999; 99US-0140823.  
29-JUN-1999; 99US-0140991.  
30-JUN-1999; 99US-0141287.  
01-JUL-1999; 99US-0141842.  
01-JUL-1999; 99US-0142154.  
02-JUL-1999; 99US-0142055.  
06-JUL-1999; 99US-0142390.  
08-JUL-1999; 99US-0142803.  
09-JUL-1999; 99US-0142920.  
12-JUL-1999; 99US-0142977.  
13-JUL-1999; 99US-0143542.  
14-JUL-1999; 99US-0143624.  
15-JUL-1999; 99US-0144005.  
16-JUL-1999; 99US-0144085.  
16-JUL-1999; 99US-0144086.  
19-JUL-1999; 99US-0144325.  
19-JUL-1999; 99US-0144331.  
19-JUL-1999; 99US-0144332.  
19-JUL-1999; 99US-0144333.  
19-JUL-1999; 99US-0144334.  
19-JUL-1999; 99US-0144335.  
20-JUL-1999; 99US-0144352.  
20-JUL-1999; 99US-0144632.  
20-JUL-1999; 99US-0144884.  
21-JUL-1999; 99US-0144814.  
21-JUL-1999; 99US-0145086.  
21-JUL-1999; 99US-0145088.  
21-JUL-1999; 99US-0145085.  
22-JUL-1999; 99US-0145087.  
22-JUL-1999; 99US-0145089.  
22-JUL-1999; 99US-0145192.  
23-JUL-1999; 99US-0145145.  
23-JUL-1999; 99US-0145218.  
23-JUL-1999; 99US-0145224.  
26-JUL-1999; 99US-0145276.  
27-JUL-1999; 99US-0145913.  
27-JUL-1999; 99US-0145918.  
27-JUL-1999; 99US-0145919.  
28-JUL-1999; 99US-0145951.  
02-AUG-1999; 99US-0146386.  
02-AUG-1999; 99US-0146388.  
02-AUG-1999; 99US-0146389.  
03-AUG-1999; 99US-0147038.  
04-AUG-1999; 99US-0147204.  
04-AUG-1999; 99US-0147302.  
05-AUG-1999; 99US-0147192.  
06-AUG-1999; 99US-0147260.  
06-AUG-1999; 99US-0147303.  
06-AUG-1999; 99US-0147416.  
09-AUG-1999; 99US-0147493.  
09-AUG-1999; 99US-0147935.  
10-AUG-1999; 99US-0148171.  
11-AUG-1999; 99US-0148319.  
12-AUG-1999; 99US-0148341.



PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149829.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	22-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143342.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157153.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145066.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145068.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149388.			

Query Match

Best Local Similarity

Matches

60;

Conservative

3;

Mismatches

7;

Indels

0;

Gaps

0;

QY

67

MDKEMPERDGVSTTKKLRMEVKSMTIVGVTSLADNEERRAFWEAGLNHCLAKPLTKDKI

126

Db

1

MDKEMPERDGVSTTKKLRMEVKSMTIVGVTSLADQEEERKAFWEAGLNHCLAKPLTKAKI

60

QY

127

IPLINQLMDA

136

Db

61

FPLISHLFDA

70

RESULT 10  
AAG21090  
ID AAG21090 standard; Protein; 66 AA.  
XX  
AC AAG21090;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23519.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 29-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR	20-AUG-1999;	99US-0149723.	ID	XX	AAG21088 standard; Protein; 71 AA.
PR	20-AUG-1999;	99US-0149829.	XX	AC	
PR	23-AUG-1999;	99US-0149902.	XX	AC	ANG21088;
PR	23-AUG-1999;	99US-0149930.	DT	XX	17-OCT-2000 (first entry)
PR	25-AUG-1999;	99US-0150566.	DE	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 23517.
PR	26-AUG-1999;	99US-0150884.	DE	XX	
PR	27-AUG-1999;	99US-0151065.	XX	XX	Protein identification; signal transduction pathway; metabolic pathway;
PR	27-AUG-1999;	99US-0151066.	XX	XX	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151080.	KW	KW	termination sequence.
PR	30-AUG-1999;	99US-0151303.	KW	KW	
PR	31-AUG-1999;	99US-0151438.	XX	OS	Arabidopsis thaliana.
PR	01-SEP-1999;	99US-0151930.	XX	XX	
PR	07-SEP-1999;	99US-0152363.	XX	XX	EP1033405-A2.
PR	10-SEP-1999;	99US-0153070.	XX	XX	06-SEP-2000.
PR	13-SEP-1999;	99US-0153758.	PD	XX	
PR	15-SEP-1999;	99US-0154018.	XX	XX	
PR	16-SEP-1999;	99US-0154039.	XX	XX	
PR	20-SEP-1999;	99US-0154779.	XX	XX	
PR	22-SEP-1999;	99US-0155139.	XX	XX	
PR	23-SEP-1999;	99US-0155486.	XX	XX	
PR	24-SEP-1999;	99US-0155659.	XX	XX	
PR	28-SEP-1999;	99US-0156458.	PR	XX	99US-0121825.
PR	29-SEP-1999;	99US-0156596.	PR	XX	99US-0123180.
PR	04-OCT-1999;	99US-0157117.	PR	XX	99US-0123548.
PR	05-OCT-1999;	99US-0157753.	PR	XX	99US-0125788.
PR	06-OCT-1999;	99US-0157865.	PR	XX	99US-0126264.
PR	07-OCT-1999;	99US-0158029.	PR	XX	99US-0126785.
PR	08-OCT-1999;	99US-0158232.	PR	XX	99US-0127462.
PR	12-OCT-1999;	99US-0158369.	PR	XX	99US-0128234.
PR	13-OCT-1999;	99US-0159293.	PR	XX	99US-0128714.
PR	13-OCT-1999;	99US-0159294.	PR	XX	99US-0129845.
PR	13-OCT-1999;	99US-0159295.	PR	XX	99US-0130077.
PR	14-OCT-1999;	99US-0159329.	PR	XX	99US-0130449.
PR	14-OCT-1999;	99US-0159330.	PR	XX	99US-0130510.
PR	14-OCT-1999;	99US-0159331.	PR	XX	99US-0130891.
PR	14-OCT-1999;	99US-0159637.	PR	XX	99US-0131449.
PR	18-OCT-1999;	99US-0159638.	PR	XX	99US-0132048.
PR	21-OCT-1999;	99US-0159584.	PR	XX	99US-0132407.
PR	21-OCT-1999;	99US-0160741.	PR	XX	99US-0132484.
PR	21-OCT-1999;	99US-0160767.	PR	XX	99US-0132485.
PR	21-OCT-1999;	99US-0160768.	PR	XX	99US-0132486.
PR	21-OCT-1999;	99US-0160770.	PR	XX	99US-0132487.
PR	21-OCT-1999;	99US-0160814.	PR	XX	99US-0132863.
PR	21-OCT-1999;	99US-0160815.	PR	XX	99US-0134256.
PR	21-OCT-1999;	99US-0160980.	PR	XX	99US-0134218.
PR	22-OCT-1999;	99US-0160981.	PR	XX	99US-0134219.
PR	22-OCT-1999;	99US-0160989.	PR	XX	99US-0134221.
PR	25-OCT-1999;	99US-0161404.	PR	XX	99US-0134370.
PR	25-OCT-1999;	99US-0161405.	PR	XX	99US-0134768.
PR	25-OCT-1999;	99US-0161406.	PR	XX	99US-0134941.
PR	26-OCT-1999;	99US-0161359.	PR	XX	99US-0135124.
PR	26-OCT-1999;	99US-0161360.	PR	XX	99US-0135353.
PR	26-OCT-1999;	99US-0161361.	PR	XX	99US-0135629.
PR	28-OCT-1999;	99US-0161920.	PR	XX	99US-0136021.
PR	28-OCT-1999;	99US-0161992.	PR	XX	99US-0136392.
PR	28-OCT-1999;	99US-0161993.	PR	XX	99US-0136782.
PR	28-OCT-1999;	99US-0161993.	PR	XX	99US-0137222.
PR	28-OCT-1999;	99US-0161993.	PR	XX	99US-0137528.
PR	28-OCT-1999;	99US-0161993.	PR	XX	99US-0137502.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0137724.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0137724.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0138094.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0138540.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0138847.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139119.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139119.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139452.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139453.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139492.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139454.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139455.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139456.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139457.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139458.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139459.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139460.
PR	28-OCT-1999;	99US-0162142.	PR	XX	99US-0139460.

Query Match 41.3%; Score 280; DB 21; Length 66;  
Best Local Similarity 84.8%; Pred. No. 1.le-23;  
Matches 56; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 71 MPERDGVSTTKLREMEVKSMTIVGVTSLADNDEEERAEAGLNHCLAKPLTKDKIPLI 130  
DB 1 MPERDGVSTTKLREMEVKSMTIVGVTSLADNDEEERAEAGLNHCLAKPLTKDKIPLI 60  
QY 131 NOLMDA 136  
DB 61 SHLFDA 66

PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144633.  
PR 20-JUL-1999; 99US-0144868.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145221.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150366.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160787.  
PR 21-OCT-1999; 99US-0160788.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 31.4%; Score 213; DB 21; Length 71;  
Best Local Similarity 73.4%; Pred. No. 3.7e-16;  
Matches 47; Conservative 3; Mismatches 8; Indels 6; Gaps 2;

QY 1 MATKSMGDIEK-----IKKKL-NVLIVDDDPNLNLIIEHKIKAIGGISOTANNGEAVII 54  
||||| |  
Db 1 MATKSTGGTEKTSIEVKKKLINVLIVDDDPNRRLHEMIIRKTTGGISQAKNGEAVIL 60  
|||||

QY 55 HRDG 58  
|||||  
Db 61 HRDG 64

RESULT 12  
AAB25159  
ID AAB25159 standard; Protein; 261 AA.  
XX  
AC AAB25159;

XX 27-NOV-2000 (first entry)  
DT Pinus radiata cell signalling involved protein SEQ ID NO:127.  
DE  
XX  
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KW environmental change; development; cell proliferation; differentiation;  
KW elongation; survival; disease resistance; nutrient metabolism.  
XX  
XX Pinus radiata.  
OS  
XX  
XX WO200042171-A1.  
PN  
XX  
XX 20-JUL-2000.  
PD  
XX  
XX 11-JAN-2000; 2000WO-US00724.  
PF  
XX  
XX 12-JAN-1999; 99US-0228986.  
XX  
PR 01-NOV-1999; 99US-0162866.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
PA  
XX Strabala TJ, Nieuwenhuizen NJ;  
PI  
XX WPT; 2000-476052/41.  
DR  
XX  
XX Isolated polynucleotide encoding a polypeptide involved in cell  
PT signaling used for generating transgenic plants with modified responses  
PT to external signals.  
XX  
XX Claim 3; Page 104-105; 527pp; English.  
PS  
XX  
XX AAA79263 to AAA79736 and AAA25100 to AAA25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
CC pine (Pinus radiata) also known as Monterey pine). The protein sequences  
CC are involved in cell signalling. The polynucleotide and protein  
CC sequences can be used to modify the response of plant cells to external  
CC signals e.g. environmental changes or pathogens during the growth and  
CC development of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to  
CC delay senescence and prolong the life of cut flowers or enhance  
CC senescence of reproductive organs to engineer sterile plants. Other  
CC modifications can be used to delay senescence in selected shelf life  
CC organs providing fruit and vegetables which have a longer shelf life  
CC between harvest and consumption, or to decrease branching frequency in  
CC forest tree species giving long stretches of valuable knot-free clear  
CC wood which can be used in solid timber furniture and veneers.  
XX  
XX Sequence 261 AA;  
SQ  
Query Match 29.8%; Score 202; DB 21; Length 261;  
Best Local Similarity 36.5%; Pred. No. 3.7e-14;  
Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;  
QY 18 VLVVDDPLNIIHEKIIKATGTSQANNCERAVIHRDGGSGSFILMDKMPERDGV 77  
DB 137 ILLVETQINRIIFGRVLQSLNLYCEAENGKVAVDYFKQ-GRTYDLVMDKEMPYNDGH 195  
QY 78 STTKLRENEVKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIPLINQ 132  
DB 196 EATROLRSMGVKPIVALTA-NTLQSKDLFEAGVDQFQSKPLSRDLVQLLDQ 249  
RESULT 13  
AAV42644  
ID AAV42644 standard; Protein; 67 AA.  
XX  
AC AAV42644;  
XX  
DT 10-JAN-2000 (first entry)

XX Brassica napus DZ2B partial fragment.  
DE  
XX  
XX Signal transduction protein; dehiscence; male sterile plant; DZ2B gene;  
KW shatter resistance; oilseed rape; response regulator protein.  
KW  
XX Brassica napus.  
OS  
XX WO9949046-A1.  
PN  
XX  
XX 30-SEP-1999.  
PD  
XX  
XX 22-MAR-1999; 99WO-GB00905.  
PF  
XX  
XX 20-MAR-1998; 98GB-0006113.  
PR  
XX (BIOG-) BIOGEMMA UK LTD.  
PA  
XX  
XX Wyatt P, Roberts JA, Whitelaw C;  
PI  
XX WPI; 1999-580449/49.  
XX  
DR N-PSDB; AA222975.  
DR  
XX A nucleic acid encoding a signal transduction protein involved in plant  
PT dehiscence, useful for producing shatter resistant male sterile plants  
PT  
XX  
XX Example 2; Fig 5; 71pp; English.  
PS  
XX The invention provides a nucleic acid encoding a signal transduction  
CC protein involved in the process of dehiscence. The nucleic acids and  
CC proteins are useful for regulating or controlling dehiscence of a pod or  
CC an anther in a plant, useful in the production of male sterile plants.  
CC The methods, etc. may be used in production of shatter resistance or  
CC shatter-delayed plants such as oilseed rape (Brassica napus). The present  
CC sequence represents the partial fragment of B. napus DZ2B protein.  
XX  
XX Sequence 67 AA;  
SQ  
Query Match 28.4%; Score 192.5; DB 20; Length 67;  
Best Local Similarity 65.7%; Pred. No. 6.7e-14;  
Matches 44; Conservative 6; Mismatches 10; Indels 7; Gaps 2;  
QY 1 MAT--KSMGDIK-----IKKLNVLIVDDPLNLIHEKIIKATGTSQANNGEAVI 53  
DB 1 MATSTSGDIKKTKSVKVKLNVLIVDDTVIRKLHENIKSIKIGGISTAKNGEAVN 60  
QY 54 IHRDGGG 60  
DB 61 IHRDGN 67  
RESULT 14  
ABG70785  
ID ABG70785 standard; Protein; 118 AA.  
XX  
XX AC ABG70785;  
XX  
XX 09-DEC-2002 (first entry)  
DT  
XX  
XX DE E. coli RscC receiver region of histidine kinase.  
XX  
XX RscC; antagonist; agonist; cytokinin receptor;  
KW receptor; signal transduction; histidine kinase; hormone; cell division;  
KW cell differentiation; agriculture; growth regulator; harvest.  
XX  
XX Escherichia coli.  
OS  
XX  
XX EP1241182-A2.  
PN  
XX  
XX 18-SEP-2002.  
PD  
XX  
XX 13-MAR-2002; 2002EP-0005749.  
PF



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2003, 19:10:32 : Search time 17 seconds  
(without alignments)  
339.487 Million cell updates/sec

Title: US-09-646-679-15  
Perfect score: 678  
Sequence: 1 MATKSMGDIKIKKLNVL.....LAKPLTKDKIPLINQLMDA 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pbp:\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pbp:\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pbp:\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pbp:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pbp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pbp:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	202	29.8	261	4	US-09-228-986-127
2	152.5	22.5	946	4	US-09-328-352-6073
3	150	22.1	1014	4	US-09-252-991A-22216
4	148	21.8	947	4	US-09-252-991A-23138
5	147	21.7	1627	4	US-09-252-991A-29359
6	145	21.4	1220	2	US-08-843-530B-36
7	144	21.2	162	4	US-09-252-991A-17557
8	139.5	20.6	1281	2	US-08-843-530B-6
9	139.5	20.6	1298	2	US-08-843-530B-2
10	139.5	20.6	1298	2	US-08-843-530B-4
11	139.5	20.6	1298	2	US-08-843-530B-34
12	139.5	20.6	1441	4	US-09-252-991A-28143
13	138.5	20.4	234	4	US-09-634-238-220
14	135.5	20.0	760	4	US-09-252-991A-25928
15	135.5	20.0	860	4	US-09-252-991A-26112
16	133.5	19.7	918	2	US-08-843-530B-35
17	131	19.3	706	4	US-09-252-991A-18607
18	131	19.3	1036	4	US-09-252-991A-27075
19	130	19.2	1017	2	US-08-843-530B-18
20	130	19.2	1181	2	US-08-843-530B-33
21	127	18.7	506	4	US-09-252-991A-18168
22	125	18.4	227	4	US-09-107-532A-6769
23	124.5	18.4	133	4	US-09-328-352-5134
24	124	18.3	971	3	US-09-112-450-2
25	124	18.3	971	4	US-09-419-291A-2
26	124	18.3	2471	3	US-09-112-450-4
27	124	18.3	2471	4	US-09-419-291A-4

28 122.5 18.1 265 4 US-09-328-352-6073 Sequence 6073, Ap  
29 122 18.0 311 4 US-09-252-991A-22932 Sequence 22932, A  
30 120.5 17.8 811 4 US-09-252-991A-22216 Sequence 22216, A  
31 119 17.6 484 4 US-09-252-991A-31677 Sequence 31677, A  
32 117 17.3 320 4 US-09-328-352-6809 Sequence 6809, Ap  
33 115.5 17.0 331 4 US-09-252-991A-23765 Sequence 23765, A  
34 115 16.7 782 4 US-09-228-986-114 Sequence 114, App  
35 113 16.7 302 4 US-09-252-991A-29564 Sequence 29564, A  
36 112.5 16.6 659 4 US-09-252-991A-16981 Sequence 16981, A  
37 112.5 16.6 129 4 US-09-252-991A-17504 Sequence 17504, A  
38 111.5 16.4 212 4 US-09-107-532A-6859 Sequence 6859, Ap  
39 111.5 16.4 234 4 US-09-328-352-5430 Sequence 5430, Ap  
40 110.5 16.3 256 4 US-09-328-986-117 Sequence 117, App  
41 110 16.2 599 4 US-09-328-986-117 Sequence 117, App  
42 109 16.1 245 4 US-09-134-001C-3779 Sequence 3779, Ap  
43 109 16.1 257 4 US-09-328-352-7337 Sequence 7337, Ap  
44 108.5 16.0 246 4 US-09-107-532A-5040 Sequence 5040, Ap  
45 108.5 16.0 250 4 US-09-634-238-221 Sequence 221, App

ALIGNMENTS

RESULT 1

US-09-228-986-127  
; Sequence 127, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 127  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-127

Query Match 29.8%; Score 202; DB 4; Length 261;  
Best Local Similarity 36.5%; Pred. No. 7e-16;  
Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;  
QY 18 VLIVDDPLNLIHEKIIKAGISQATNGEEAVIHRDGGSSFDILMDKEMPERDGV 77  
Db 137 ILLVEDTQINRIIFGRVLSLNLYCEEAEKGVAVDFKQ-GRTYDVLMDKEMPVMDGH 195  
QY 78 STTKKIREMEVKSIMVGVTSADNEEERRAEAGLNHCLAKPLTKKIIPLNQ 132  
Db 196 EATQLRSGVKTPIVALTA-NTLOSKDLTFEAGVDDFQSKPLSRDLVQLLQ 249

RESULT 2

US-09-328-352-7973  
; Sequence 7973, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7973  
; LENGTH: 946  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7973

```
Query Match      22.5%; Score 152.5; DB 4; Length 946;
Best Local Similarity 30.2%; Pred. No. 3.4e-09;
Matches 39; Conservative 34; Mismatches 41; Indels 15; Gaps 4;

QY 16 LNLVIVDDDDPLNLIHKEIKIAGGISQTANNGEAVIIHRDGGSS----FDLILMDKEM 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 674 LHILAVDDHLPNLIVLEALLGELNVKTKALSGQEAALNIQRIQDKLKPFDLVFMDIQM 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 72 PERDGVSTTKKLEEM-----EVKSMIVGVT--SLADNEEERAFMEAGLNHCLAKPLTK 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 734 PVMSGIDTTRAIRSLESLDGEMLPIALTAHALAD---EKQKLVGMNDYVTKPIQM 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 124 DKIIPLNQ 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 791 EQIIQILTQ 799
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-09-252-991A-31338
; Sequence 31338, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31338
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31338

Query Match      22.1%; Score 150; DB 4; Length 1014;
Best Local Similarity 25.5%; Pred. No. 7.4e-09;
Matches 35; Conservative 36; Mismatches 56; Indels 10; Gaps 4;

QY 2 ATKSMGDIKIK--KKNLVIVDDDDPLNLIHKEIKIAGGISQTANNGEAVIIHRDGG 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 868 ARSGVAEVEERKEARALSILLAEHPFNRLTLTMQLESGLHRVSTEDGEEA--FERWQ 925
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 60 SSFDLILMDKEMPERDGVSTTKKLEMEV-----KSMIVGVTSLADNEEERAFMEAGLN 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 926 EDFDVVITDGMFPRMDGYELARRISQEAALGRRRCILVIALTASAEKDALERC-LAAGMD 984
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 115 HCLAKPLTKDKIPLIN 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 985 RVLFKPTLDELARLN 1001
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-09-252-991A-29359
; Sequence 29359, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 29359
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29359

Query Match      21.8%; Score 148; DB 4; Length 947;
Best Local Similarity 28.8%; Pred. No. 1.1e-08;
Matches 34; Conservative 33; Mismatches 43; Indels 8; Gaps 4;

QY 18 LVIVDDDDPLNLIHKEIKIAGGISQTANNGEAV--IHRDGGSSFDLILMDKEMPERDG 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 689 LLCVDNDPANLLVQLLSDLGAQVTAVDGSAALEVQRE---RFDLVFMDVQPMGMDG 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 77 VSTTKKLEMEVKSMT--VGVTSIADN--EEERRAFMEAGLNHCLAKPLTKDKIPLI 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 746 RQATEAIRWEAEREVSPVPIALTAAHLSNEKRALQMGDDYLTKTIDEQOLAQVV 803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-252-991A-20395
; Sequence 20395, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20395
; LENGTH: 1627
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20395

Query Match      21.7%; Score 147; DB 4; Length 1627;
Best Local Similarity 31.9%; Pred. No. 3.3e-08;
Matches 37; Conservative 26; Mismatches 45; Indels 8; Gaps 3;

QY 14 KKNLVIVDDDDPLNLIHKEIKIAGGISQTANNGEAVIIHRDGGSSFDLILMDKEMPE 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1371 RRLQVLVVDHDAVNRQLHQQLSFLGHVDEAEAGLSALNLH--GQPFDMVITDCHMPL 1428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 74 RDGVSTTKKLEMEVKS-----MIVGVTSLADNEEERAFMEAGLNHCLAKPLTKD 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1429 MSGSLARSIRQEEENGEPVWIIGLTADAQAPPEIERC-IOAGMNECLIKPIGLD 1483
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-08-843-530B-36
; Sequence 36, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,530B  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: UTC-02717  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1220 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-843-530B-36

Query Match 21.4%; Score 145; DB 2; Length 1220;  
Best Local Similarity 29.0%; Pred. No. 3.8e-08;  
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DIEIKKKLNVLIVDDPLNLIHEKIKA---IGGIS--QTANNGEEA---VVIHRDGG 59  
DB 1080 DDDKNEISVKILVVEDNHVN---QEVIKRMLNLEGIEIETACDQGEAFDKVKELTSKG 1135  
QY 60 SFEDLILMDKEMPERDGVSTTKKL-REMEVKSMIVGVTSLADNEERAFMEAGLNHCLA 118  
DB 1136 ENYNIEMDVQMPVQGLLSTKMI-RRLDGYTSPVIALTAFAADDSNKEC-LESGMNGFLS 1194

QY 119 KPLTKDKIPLINOLMDA 136  
DB 1195 KPIKPKLTKLTBECAA 1212

RESULT 7  
US-09-252-991A-17557  
Sequence 17557, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17557  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17557

Query Match 21.2%; Score 144; DB 4; Length 162;  
Best Local Similarity 30.9%; Pred. No. 2.6e-09;  
Matches 38; Conservative 28; Mismatches 49; Indels 8; Gaps 3;

QY 18 VLVVDDPLNLIHEKIKAIGGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERDGV 77  
DB 35 ILVAEDNPVNLVVRGFLAKRGYAVRLAGNGLADEYLRDPNGIQLILMDGEMPMDFG 94  
QY 78 STTKKLREME-----VKSMIVGVTSLADNEERAFMEAGLNHCLAKPLTKDKIIFLIN 131

Db 95 EATRLIRREERAQCWPRVPVIALTAHLD--BHRRAGIEAGMDAYLGKPVDRAEIATLE 152  
QY 132 QLM 134  
Db 153 RLL 155

RESULT 8  
US-08-843-530B-6  
Sequence 6, Application US/08843530B  
Patent No. 5939306  
GENERAL INFORMATION:  
APPLICANT: Sellitrennikoff, Claude  
APPLICANT: Agnan, Jacqueline  
APPLICANT: Alex, Lisa A.  
APPLICANT: Simon, Melvin I.  
TITLE OF INVENTION: Osmosensing Histidine Kinases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,530B  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: UTC-02717  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1281 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-843-530B-6

Query Match 20.6%; Score 139.5; DB 2; Length 1281;  
Best Local Similarity 31.1%; Pred. No. 1.8e-07;  
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY 2 ATKSMGDIEKIKKLNVLIVDDPLNLIHEKIKAIGGISQTANNGEEAV-IHRDGG 60  
DB 1075 ATPSLAD---NTKSFELLAEADNVNQLAVKILEKHHVTVVVGNGEEAVEAVKR---K 1128  
QY 61 SFDLILMDKEMPERDGVSTTKKLREME-----VKSMIVGVTSLADNEERAFMEAGLNH 115  
DB 1129 KFDVILMDVQMPIMGGFETAKIREYERSLGSRTPPIALTATAHA-MMGDREKCIQAO 1187

QY 116 CLAKPLTKDKIPLI 130  
DB 1188 YLSKPLQONHLIQT 1202

RESULT 9  
US-08-843-530B-2  
Sequence 2, Application US/08843530B  
Patent No. 5939306  
GENERAL INFORMATION:  
APPLICANT: Sellitrennikoff, Claude

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-843-530B-4

Query Match          20.6%; Score 139.5; DB 2; Length 1298;
Best Local Similarity 31.1%; Pred. No. 1.8e-07;
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY      2 ATKSMGDIETKKKLNLIVDDPLNLIHEKIIKAIGGISOTANNGEAV-IIHROGG 60
        || || | : | : | : | : | : | : | : | : | : | : | : | : |
Db      1075 APPSLAD---NTRKEILLAEEDNTVQRALVKLEYHVVTVNGEAVEAYKR--K 1128
QY      61 SPDLILDKENPERDGYSTTKKREME-----YKSMIVGVTSIADNEEREAFWEAGLNH 115
        ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db      1129 KFDVLMDVQPMGPMPGFEATAKIREVERSLGSORTPIALTAAH-MMGDRCKICQAOMDE 1187
QY      116 CLAKPLTKDKIPLI 130
        ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db      1188 YLSKPLQNHLIQTI 1202
        ||||| : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 11
US-08-843-530B-34
Sequence 34, Application US/08843530B
Patent No. 5939306
GENERAL INFORMATION:
APPLICANT: Selltreinnikoff, Claude
APPLICANT: Agnan, Jacqueline
APPLICANT: Alex, Lisa A.
APPLICANT: Simon, Melvin I.
TITLE OF INVENTION: Osmosensing Histidine Kinases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELECOMMUNICATION INFORMATION:

```

```
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1298 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-843-530B-34

Query Match          20.6%; Score 139.5; DB 2; Length 1298;
Best Local Similarity 31.1%; Pred. No. 1.8e-07;
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

Qy 2 ATKSMGDIKIKKLVNLIIVDDPLNLIHKEIKIAGIGISOTANNGEAV-IHRDGS 60
Db 1075 ATPLAD---NTKSFELLAEEDTNVORLAVKILEXYHHVTVVGVNGEAVKVR---K 1128
Qy 61 SFDLILMDKEMPERDGVSTTKKLREME-----VKSMIVGVTSLADNEEERAFMEAGLNH 115
Db 1129 KFDVILMDVQPMGGEATAKIREYERSIGSORTPIIALTAHA-MMGDREKCIQAQND 1187
Qy 116 CLAKPLTKDKIPLI 130
Db 1188 YLSKPLQQNHLIQTI 1202

RESULT 12
US-09-252-991A-28143
; Sequence 28143, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28143
; LENGTH: 1441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28143

Query Match          20.6%; Score 139.5; DB 4; Length 1441;
Best Local Similarity 28.8%; Pred. No. 2.2e-07;
Matches 34; Conservative 30; Mismatches 43; Indels 11; Gaps 3;

Qy 15 LNVLIIVDDPLNLIHKEIKIAGIGISOTANNGEAVIHRDGGSSFDLILMDKEMPERD 75
Db 1182 VRLIIVDDNLRQVAEALLSSEGARVALADGGLAGVQQVLEASVPFDVAVLMDQMPPDID 1241
Qy 76 GVSVTTKLRE-----NEVKSMTIVGVTSLADNEEERAFMEAGLNHCLAKPLTKDKII 127
Db 1242 GLEATRRIRADGRAGIPLANTAN-ASLAD----REACLAGMNDHVAKPIDKELV 1294

RESULT 13
US-09-634-238-220
; Sequence 220, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
```

```
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.104301
; CURRENT APPLICATION NUMBER: US/09/634.238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-220

Query Match          20.4%; Score 138.5; DB 4; Length 234;
Best Local Similarity 30.2%; Pred. No. 2e-08;
Matches 35; Conservative 31; Mismatches 47; Indels 3; Gaps 2;

Qy 18 VLIVDDPLNLIHKEIKIAGIGISOTANNGEAVIHRDGGSSFDLILMDKEMPERDGV 77
Db 4 ILIVDDPEAITLLOYNLEAEHYQVETATDGEA--LDKVRSEPFDFILDLMLFSLSL 61
Qy 78 STTKKLREMEVKSMTIVGVTSLADNEEERAFMEAGLNHCLAKPLTKDKIPLINOL 133
Db 62 DVTCKIREEKIQTPIMILTA-KDNETDKIVLGELGADDYVTKPFPREIARIKAI 116

RESULT 14
US-09-252-991A-25928
; Sequence 25928, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25928
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25928

Query Match          20.0%; Score 135.5; DB 4; Length 760;
Best Local Similarity 28.0%; Pred. No. 2.5e-07;
Matches 35; Conservative 31; Mismatches 54; Indels 5; Gaps 3;

Qy 4 KSMGDIKIKKLVNLIIVDDPLNLIHKEIKIAGIGISOTANNGEAVIHRDGGSSFD 63
Db 612 QAAGDAAPVAAGQBIILVEDPNVQTVIEAMLSLGYRTVLVADGIQAV--RSAERQRYD 669
Qy 64 LILMDKEMPERDGVSTTKKLREMEVKSMTIVGVTSLADN--EEERRAFMEAGLNHCLAKPL 121
Db 670 AILMDCLPLVDGYSATREIRAQE-NGRQVPIIALTANALQDRENCIQAQMNDYLAKEPF 728

Qy 122 TKDKI 126
Db 729 KRAEL 733

RESULT 15
US-09-252-991A-26112
; Sequence 26112, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26112
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26112

Query Match      20.0%; Score 135.5; DB 4; Length 860;
Best Local Similarity 26.5%; Pred. No. 3e-07;
Matches 36; Conservative 35; Mismatches 56; Indels 9; Gaps 3;

QY      1 MATKSGDTEKIKKLNVLIVDDPLNLIHEKIIKAIGISQTANNGEEAVIIHRDGS 60
Db      722 IAPSSASDTTGEQRNTRVLLVEDNPVNLVAKGLLHLKGCQVIAEHGLNALKMLEE--H 779

QY      61 SFDLILMDKEMPERDGVSTTKLREM----EVKSMIVGVTSLADNEEERRAFMEAGLNHC 116
Db      780 PIDLVLMDCNMPYMDGYEATQIRDSGRWGGLPIIALTANALPDERERCRA---AGMDDY 836

QY      117 LAKPLTKDKTIPLINQ 132
Db      837 LAKPFHDELKALDR 852

Search completed: August 13, 2003, 19:13:02
Job time : 18 secs
```





```
/ APPLICANT: Madden, Kevin
/ APPLICANT: Maxon, Mary
/ APPLICANT: Milne, Todd
/ APPLICANT: No. US20020128250Alman, Thea
/ APPLICANT: Royer, John
/ APPLICANT: Salama, Sofie
/ APPLICANT: Sherman, Amir
/ APPLICANT: Silva, Jeff
/ APPLICANT: Summers, Eric
/ TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
/ FILE REFERENCE: 109272.147
/ CURRENT APPLICATION NUMBER: US/09/801.368
/ CURRENT FILING DATE: 2001-03-07
/ PRIOR APPLICATION NUMBER: US 09/487,558
/ PRIOR FILING DATE: 2000-01-19
/ PRIOR APPLICATION NUMBER: US 60/160,587
/ PRIOR FILING DATE: 1999-10-20
/ NUMBER OF SEQ ID NOS: 440
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 332
/ LENGTH: 1220
/ TYPE: PRT
/ ORGANISM: Saccharomyces cerevisiae
US-09-801-368-332

Query Match 21.4%; Score 145; DB 10; Length 1220;
Best Local Similarity 29.0%; Pred. No. 6.6e-07;
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DIEIKKLNVLIVDDPLNLIIEHKIKA--IGGIS--QTANNGEA--VHHRDGG 59
Db 1080 DDKNETSKVLIVEDHNVN---QEVKRMNLGEGENIELACDGEADPKVELTSKG 1135
QY 60 SSFDILMDKEMPERDGVSTTKKL-REMEVKSMTIVGVTSLADNEERARRAMEAGLNHCLA 118
Db 1136 ENYNMFDMVQMPKVDGELLSTKMIRRLDGYTSPVALTAFADDSNKEC-LESGMNGFLS 1194
QY 119 KPLTKDKIPLINQLMDA 136
Db 1195 KPDKPRKLTILTEFCAA 1212

RESULT 6
US-09-801-368-366
/ Sequence 366, Application US/09801368
/ Patent No. US20020128250A1
/ GENERAL INFORMATION:
/ APPLICANT: Busby, Robert
/ APPLICANT: Cali, Brian
/ APPLICANT: Hecht, Peter
/ APPLICANT: Holtzman, Doug
/ APPLICANT: Madden, Kevin
/ APPLICANT: Maxon, Mary
/ APPLICANT: Milne, Todd
/ APPLICANT: No. US20020128250Alman, Thea
/ APPLICANT: Royer, John
/ APPLICANT: Salama, Sofie
/ APPLICANT: Sherman, Amir
/ APPLICANT: Silva, Jeff
/ APPLICANT: Summers, Eric
/ TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
/ FILE REFERENCE: 109272.147
/ CURRENT APPLICATION NUMBER: US/09/801.368
/ CURRENT FILING DATE: 2001-03-07
/ PRIOR APPLICATION NUMBER: US 09/487,558
/ PRIOR FILING DATE: 2000-01-19
/ PRIOR APPLICATION NUMBER: US 60/160,587
/ PRIOR FILING DATE: 1999-10-20
/ NUMBER OF SEQ ID NOS: 440
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 366
/ LENGTH: 712
/ TYPE: PRT

/ ORGANISM: Saccharomyces cerevisiae
US-09-801-368-366

Query Match 21.2%; Score 143.5; DB 10; Length 712;
Best Local Similarity 27.9%; Pred. No. 4.5e-07;
Matches 39; Conservative 28; Mismatches 42; Indels 31; Gaps 3;

QY 10 EKIKKLNVLIVDDPLNLIIEHKIKAIGGISQTANNGEAAVHHRDGGSSFDILMDK 69
Db 498 EKVEPKINVLIVEDNVIQNSFLRKHKISKYKLAKNQEAQVNIWKEGG--LHIFMDL 555
QY 70 EMPERDGVSTTKKLREMEVKS-----MIVGVTSLADN 101
Db 556 OLPVLSGIEAAKQIRDEKQNGIGIOKSLNHSNLEKTSKRFSOAPVYIIVALTASNSQ 615
QY 102 EEERRAFMEAGLNHCLAKPL 121
Db 616 MDKRKALL-SGCNDYLTQPV 634

RESULT 7
US-10-126-120-2
/ Sequence 2, Application US/10126120
/ Publication No. US20030108526A1
/ GENERAL INFORMATION:
/ APPLICANT: Sakakibara, Hitoshi
/ APPLICANT: Takei, Kentaro
/ TITLE OF INVENTION: MICROORGANISMS FOR USE IN THE MEASUREMENT OF ENVIRONMENTAL FAC
/ FILE REFERENCE: 11127-004001
/ CURRENT APPLICATION NUMBER: US/10/126.120
/ CURRENT FILING DATE: 2002-04-19
/ PRIOR APPLICATION NUMBER: JP 2001-291059
/ PRIOR FILING DATE: 2001-09-25
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 974
/ TYPE: PRT
/ ORGANISM: Zea mays
US-10-126-120-2

Query Match 20.3%; Score 137.5; DB 15; Length 974;
Best Local Similarity 24.6%; Pred. No. 3.2e-06;
Matches 35; Conservative 32; Mismatches 46; Indels 29; Gaps 3;

QY 17 NVLIVDDPLNLIIEHKIKAIGGISQTANNGEAAVHHRDGGSSFDILMDKEMP 72
Db 830 NILVDDNKVNLRVAAALKKYGANVSCVESGKDAISLQPPHR-----EDACFMDVQMP 884
QY 73 ERDGVSTTKKLREMEVKSMTIVGVTSLADNEERAA-----EWEAG 112
Db 885 ENDGFEATGQIRQMELKANEERKNKLASIEGSTTAHYHLPVLAMTADVIQATYEECIKSG 944
QY 113 LNHCLAKPLTKDKIPLINQLM 134
Db 945 MDGYVSKPDEEQLYQAVSRLV 966

RESULT 8
US-09-918-508-7
/ Sequence 7, Application US/09918508
/ Patent No. US20020177162A1
/ GENERAL INFORMATION:
/ APPLICANT: KAKIMOTO, TATSUO
/ APPLICANT: HIGUCHI, MASAYUKI
/ APPLICANT: INOUE, TSUTOMU
/ TITLE OF INVENTION: ANALYSIS OF AGONIST-ACTIVITY AND ANTAGONIST-ACTIVITY
/ TITLE OF INVENTION: TO CYTOKININ RECEPTOR
/ FILE REFERENCE: Q65478
/ CURRENT APPLICATION NUMBER: US/09/918.508
/ CURRENT FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: JP 2001-073812
/ PRIOR FILING DATE: 2001-03-15
```



```
RESULT 12
US-09-815-242-11131
; Sequence 11131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11131
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11131

Query Match      18.7%  Score 126.5;  DB 9;  Length 227;
Best Local Similarity 29.1%;  Pred. No. 6.7e-06;
Matches 34;  Conservative 30;  Mismatches 48;  Indels 5;  Gaps 2;

QY 18 VLVDDPLNLIHEKIIKAIGGISQTANNNGEEAVIIHRDGGSSFDLIIMDKEMPERGV 77
Db 4 LLLVDDDIETELSTLLELEGFVETANNGLAAL---QKLNESYKLVLLDVMKPLNGI 60
QY 78 STTKKLRMEYKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKDKIPLINOLM 134
Db 61 ETLKEIR--KVSNNPVMMLTARGEDIDRVGLGLEIGADDCLPKPFENDRELIAIKAIL 115

RESULT 13
US-10-156-761-12532
; Sequence 12532, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9975
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9975

Query Match      18.5%  Score 125.5;  DB 15;  Length 227;
Best Local Similarity 29.2%;  Pred. No. 8.6e-06;
Matches 38;  Conservative 28;  Mismatches 53;  Indels 11;  Gaps 5;

QY 10 EKIKKLNVLIVDDPLNLIHEKIIKAIG---GIS--QTANNNGEEAVIIHRDGGSSFDL 64
Db 4 EESRKPAPVVVADD---QTVVREGIVMLLGLLPGLIEYVGAAGDGHEAVKLVAE--LNPDV 58
QY 65 ILMDKEMPERDGVSTTKKLRMEYKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKD 124
Db 59 VMDLRMPRCDCVEATRRIRAEHPGTOVVVLTYYADDESIFPA--LRAGARGYLTKDAGD 117
QY 125 KIIPLINOLM 134
Db 118 EIVRAVESVL 127

RESULT 15
US-09-815-242-13783
; Sequence 13783, Application US/09815242
```

```
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12532
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12532

Query Match      18.6%  Score 126;  DB 15;  Length 203;
Best Local Similarity 29.5%;  Pred. No. 6.5e-06;
Matches 36;  Conservative 20;  Mismatches 46;  Indels 20;  Gaps 3;

QY 17 NVLIVDDPLNLIHEKIIKAIGGISQ--TANNNGEEAVIIHRDGGSSFDLIIMDKEMPER 74
Db 3 SVLVCDSDPLAREALRRVATVPGVERVTAAANGEE--VLRWGDADRSDLILMDVMPGL 60
QY 75 DGVSTTKKLRMEYKSMIVGVTSLAD-----NEEERRAFMEAGLNHCLIA 118
Db 61 GGVTVRLLSADPGARIIMLTVAEDLDGVALAAGARGYLHKDSRAELRATVYQALA 120
QY 119 KP 120
Db 121 DP 122

RESULT 14
US-10-156-761-9975
; Sequence 9975, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9975
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9975

Query Match      18.5%  Score 125.5;  DB 15;  Length 227;
Best Local Similarity 29.2%;  Pred. No. 8.6e-06;
Matches 38;  Conservative 28;  Mismatches 53;  Indels 11;  Gaps 5;

QY 10 EKIKKLNVLIVDDPLNLIHEKIIKAIG---GIS--QTANNNGEEAVIIHRDGGSSFDL 64
Db 4 EESRKPAPVVVADD---QTVVREGIVMLLGLLPGLIEYVGAAGDGHEAVKLVAE--LNPDV 58
QY 65 ILMDKEMPERDGVSTTKKLRMEYKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKD 124
Db 59 VMDLRMPRCDCVEATRRIRAEHPGTOVVVLTYYADDESIFPA--LRAGARGYLTKDAGD 117
QY 125 KIIPLINOLM 134
Db 118 EIVRAVESVL 127

RESULT 15
US-09-815-242-13783
; Sequence 13783, Application US/09815242
```

1



A:Accession: G65010  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1197 <BLAT>  
A:Cross-references: GB:AE000325; GB:U00096; NID:gl789709; PIDN:AAC75429.1; PID:gl788713  
A:Experimental source: strain K-12, substrain MG1655  
R:Utsumi, R.  
submitted to JIPID, January 1993  
A:Description: Newly identified genes involved in signal transduction of Escherichia coli  
A:Reference number: JU0220  
A:Accession: JU0221  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-151, 'P', 153-241, 'PL', 244-274, 'R', 276-419, 'FE', 422-738, 'D', 740-757, 'K', 759-  
R:Utsumi, R.; Katayama, S.; Iinaguchi, M.; Horie, T.; Ikeda, M.; Igaki, S.; Nakagawa, H.  
Gene 140, 73-77, 1994  
A:Title: Newly identified genes involved in the signal transduction of Escherichia coli  
A:Reference number: I41198; MUID:94171083; PMID:8125343  
A:Accession: I41200  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-151, 'P', 153-241, 'PL', 244-274, 'R', 276-419, 'FE', 422-738, 'D', 740-757, 'K', 759-  
A:Cross-references: GB:D14008; NID:g456162; PIDN:BAA03108.1; PID:g216554  
C:Genetics:  
A:Gene: evgS  
C:Superfamily: evgS protein; response regulator homology  
C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase;  
F:687-935/Domain: sensor histidine kinase homology <SHK>  
F:961-1070/Domain: response regulator homology <RRH>  
F:721/Binding site: phosphate (His) (covalent) #status predicted  
F:1009/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 26.3%; Score 178.5; DB 1; Length 1197;  
Best Local Similarity 34.4%; Pred No 5 le-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;

Qy 12 IKKKLVLIIVDDPLNIITHEKIKAIGGISQTANNGEAVIITHRGSGSFLLIMDKEM 71  
:||||| | : : : : : ||||| :||| :|  
Db 955 LPKLSLIIADDPTRNLRLKQLNLGYDYVDATGVQA--LHKVSMQHLYDLITDVNM 1012  
:||||| | : : : : : ||||| :||| :|  
Qy 72 PERDGVSTTKKLMEVKMSIVGTVSLADNEERAFMEAGLNHCLAKPLTKDIPLIN 131  
:||||| | : : : : : ||||| :||| :|  
Db 1013 PNMDGFELTRKLEQNSSLPIWGLTANA-QANEREKGLSCGMNLCFLKPLTLVDVLTHLS 1071  
:||||| | : : : : : ||||| :||| :|  
Qy 132 QL 133  
||  
Db 1072 QL 1073

RESULT 3  
A91035  
probable sensor for regulator EvgA [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 24-Aug-2001  
C:Accession: A91035  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A91035  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1197 <BLAT>  
A:Cross-references: GB:BA000007; PIDN:BA836672.1; PID:gl3362719; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS3249  
C:Superfamily: evgS protein; response regulator homology

Query Match 26.0%; Score 176.5; DB 2; Length 1197; GC 40.1%;  
Best Local Similarity 34.4%; Pred. No. 7.5e-07;  
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;

QY 12 IKKKINLVIVDDPDLNLIITHEKTIKRAIGGIGISQTANNGEBAVITHRDGGSSFFLIIMDKREM 71  
 Db 955 LPKLSILITADDPTRNRLIKLRQLNLLGYDVEADGVQA--LHKVSMCHYDLLITDYNM 1012  
 QY 72 PERDGVSTTKKLRGMEVSKMIVCVTSIADNEEBERRAFMEAGLNHCLAKPLTVDKIKIPLIN 131  
 Db 1013 PNYDGFELTRKLRQNSLPIWGLTANA-QANERKGLNCGMNLCLFKPLTLDVLKTHLS 1071  
 QY 132 QL 133  
 Db 1072 QL 1073  
 RESULT 4  
 C85879  
 probable sensor for regulator EvgA evgS [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: C85879  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoculis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: C85879  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1197 <STO>  
A:Cross-references: GB:AE005174; NID:gl2516740; PIDN:AAG57495.1; GSPDB:GN00145; UWGFP:  
A:Experimental source: strain C157:H7, substrain EDL933

A: Gene: evgS									
C: Superfamily: evgS protein; response regulator homology									
Query Match	26.0%	Score	176.5;	DB	2;	Length	1197;		
Best Local Similarity	34.4%;	Pred. No.	7.5e-07;						
Matches	42;	Conservative	25;	Mismatches	52;	Indels	3;	Gaps	2;
QY	12	IKKLNVLIVDDPLNLI	IIHEKIIKAGTGSQTANNGEEAVII	HRGGSSFDLIMDKEM	71				
		: : : : : : : : : :	: : : : : : : : : :	: : : : : : : : : :					
DB	955	LPEKLSIIADDDHPNLRLLK	ROLNLGLGYDVEATDGQA--	LHKYSMQHYDLSITDVNM	1012				
		: : : : : : : : : :	: : : : : : : : : :	: : : : : : : : : :					
QY	72	PERGVSTTKKLREMEVKS	MTVGTSLADNEEERAPMEAGLNH	CLAKPLTKDKIIPLN	131				
		: : : : : : : : : :	: : : : : : : : : :	: : : : : : : : : :					
DB	1013	PNVDGFELTKLRQNSSLF	INGLTANA-QANEREKGLNCGM	LCLEKPLTLDVLKTHLS	1071				
		: : : : : : : : : :	: : : : : : : : : :	: : : : : : : : : :					
QY	132	QL	133						
		: :							
DB	1072	OL	1073						

RESULT 5  
H87640  
sensor histidine kinase/response regulator [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87640  
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Odson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kocot, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: H87640  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <STO>  
A:Cross-references: GB:A0005673; NID:g13424832; PID:AAK25124.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3162

Query Match  
25.7%; · Score 174; DB 2; Length 394;

```
Best Local Similarity 31.7%; Pred. No. 3.4e-07;
Matches 39; Conservative 31; Mismatches 49; Indels 4; Gaps 2;

QY 14 KKLNVLVDDPLNLIIEHKIKAIGGISOANNGEAVIIHRDGGSSFDLILMDKEMPE 73
Db 262 RGAHLLIVDDNATNRVAFALCMFECSEGVADGVEAVEMARG--RFDLILMDIKMPR 319
QY 74 RGVSTTKLREMEYKSMIVGTVSLADNEE--ERRAFMEAGLNHCLAKPLTKDKI 131
Db 320 MDGVAATRAIRELSGRSSAAPVALTANADPADVHTYLAAGQDVKERPKERLALVLN 379
QY 132 QLM 134
Db 380 SLL 382

RESULT 6
AD0790
Sensor protein RscC (EC 2.7.3.-) [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0790
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0790
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-948 <FAR>
A:Cross-references: GB:AL513382; PIDN:CAD07502.1; PID:gl6503497; GSPDB:GN00176
C:Genetics:
A:Gene: rscC
A:Superfamily: rscC protein; response regulator homology
C:Keywords: phosphotransferase

Query Match 25.6%; Score 173.5; DB 2; Length 948;
Best Local Similarity 33.9%; Pred. No. 1e-06;
Matches 37; Conservative 33; Mismatches 36; Indels 3; Gaps 2;

QY 18 VLIVDDPLNLIIEHKIKAIGGISOANNGEAVIIHRDGGSSFDLILMDKEMPERDGV 77
Db 827 ILVVDHPINRLLADQLGSLGYQCKTANDGVDAENVL--SKNATDIVLSDVNPMDGY 884
QY 78 STTKKLREMEYKSMIVGTVSLADNEEERRAFMEAGLNHCLAKPLTKDKI 126
Db 885 RLQTRIQRLGLTLPVIGVTANALAEKQRC-LESQMDSCLSKPVTLDAL 932

RESULT 7
BVECCC
sensor protein rscC (EC 2.7.3.-) - Escherichia coli (strain K-12)
N:Alternate names: regulatory protein rscC
C:Species: Escherichia coli
C:Date: 30-Jun-1991 #sequence_revision 17-Oct-1997 #text_change 01-Mar-2002
C:Accession: H64991; JVO069; A48659
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64991
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-933 <BIAT>
A:Cross-references: GB:AE00311; GB:U00096; NID:gl788547; PIDN:AA075278.1; PID:gl788548;
A:Experimental source: strain K-12, substrain MG1655
R:Stout, V.; Gottesman, S.
J. Bacteriol. 172, 659-669, 1990

A:Title: RcsB and RcsC: a two-component regulator of capsule synthesis in Escherichia
A:Reference number: JVO069; MUID:90130299; PMID:2404948
A:Accession: JVO069
A:Molecule type: DNA
A:Residues: 1-112, 'IG', 115-918, 'S', 920-933 <STO>
A:Cross-references: GB:M28242; NID:gl47524; PIDN:AAA24503.1; PID:gl147525
A:Experimental source: strain K12
R:Jayaratne, P.; Keenleyside, W.J.; MacLachlan, P.R.; Dodgson, C.; Whitfield, C.
J. Bacteriol. 175, 5384-5394, 1993
A:Title: Characterization of rcsB and rcsC from Escherichia coli O9: K30:H12 and exar
A:Reference number: A48659; MUID:93374832; PMID:8366025
A:Accession: A48659
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-112, 'IG', 115-258, 'V', 300-759, 'E', 761-873, 'T', 875-921, 'V', 923-931, 'E', 9
A:Cross-references: GB:LI1272; NID:gl47527; PIDN:AAA24505.1; PID:gl147528
A:Experimental source: strain K30 (O9:K30:H12)
C:Comment: This protein acts as the sensor of the two-component regulatory system to
C:Genetics:
A:Gene: rscC
A:Map position: 48 min
A:Superfamily: rscC protein; response regulator homology
C:Keywords: autophosphorylation; capsule synthesis; phosphohistidine; phosphoprotein.
F:4-26/Domain: transmembrane #status predicted <TM1>
F:298-318/Domain: transmembrane #status predicted <TM2>
F:811-920/Domain: response regulator homology <RRH>
F:463/Binding site: phosphate (His) (covalent) #status predicted
F:859/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 25.3%; Score 171.5; DB 1; Length 933;
Best Local Similarity 33.9%; Pred. No. 1.5e-06;
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

QY 18 VLIVDDPLNLIIEHKIKAIGGISOANNGEAVIIHRDGGSSFDLILMDKEMPERDGV 77
Db 811 ILVVDHPINRLLADQLGSLGYQCKTANDGVDAENVL--SKNHDIVLSDVNPMDGY 868
QY 78 STTKKLREMEYKSMIVGTVSLADNEEERRAFMEAGLNHCLAKPLTKDKI 126
Db 869 RLQTRIQRLGLTLPVIGVTANALAEKQRC-LESQMDSCLSKPVTLDVI 916

RESULT 8
C91017
sensor for ctr capsule biosynthesis [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C91017
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, (
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11238796
A:Accession: C91017
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-933 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036530.1; PID:gl3362576; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS107
C:Superfamily: rscC protein; response regulator homology

Query Match 25.3%; Score 171.5; DB 2; Length 933;
Best Local Similarity 33.9%; Pred. No. 1.5e-06;
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

QY 18 VLIVDDPLNLIIEHKIKAIGGISOANNGEAVIIHRDGGSSFDLILMDKEMPERDGV 77
Db 811 ILVVDHPINRLLADQLGSLGYQCKTANDGVDAENVL--SKNHDIVLSDVNPMDGY 868
QY 78 STTKKLREMEYKSMIVGTVSLADNEEERRAFMEAGLNHCLAKPLTKDKI 126
```

Db 869 RLTORIRQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 9

ES5861

Hypothetical protein rscC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: E85861

R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 523-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85861

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 <STO>

A:Cross-references: GB:AE005174; NID:g12516556; PIDN:AAG57353.1; GSPDB:GN00145; UMGF:234

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

C:Superfamily: rscC protein; response regulator homology

Query Match 25.3%; Score 171.5; DB 2; Length 933;

Best Local Similarity 33.9%; Pred. No. 1.5e-06;

Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

QY 18 VLIVDDPLNLIIEKIIKAIGGISTQANNGEAVIIHRDGGSFLLIMDKEMPERDGV 77

DB 811 ILVDDHPINRLRLADQLGSGYCKTANDGVDAINVL--SKNHIDIVLSDVNNPNMDGY 868

QY 78 STTKKLEMEVKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126

DB 869 RLTORIRQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 10

AD0149

Two component sensor kinase/response regulator protein RscC (EC 2.7.3.-) [imported] - Yersinia pestis

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD0149

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0149

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-957 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90055.1; PID:g15979276; GSPDB:GN00175

A:Gene: rscC

C:Genetics:

C:Superfamily: rscC protein; response regulator homology

C:Keywords: phosphotransferase

Query Match 25.3%; Score 171.5; DB 2; Length 957;

Best Local Similarity 35.1%; Pred. No. 1.5e-06;

Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;

QY 16 LNVIVDDPLNLIIEKIIKAIGGISTQANNGEAVIIHRDGGSFLLIMDKEMPERD 75

DB 835 LQILVDDHPINRLRLADQLTILGVRVITANDGLDALVALNT--NTVDMVLTVNPNMD 892

QY 76 GVSTTKKLEMEVKSMIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126

DB 893 GYRLTERLQNLNHPFIIGVTANALAEKQRC-IEAGMDCLSKPVTLDTL 942

RESULT 11

ES7644

Sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: E87644

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Lub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, n.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <STO>

A:Cross-references: GB:AE005673; NID:g13424865; PIDN:AAK25153.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3191

Query Match 24.0%; Score 162.5; DB 2; Length 642;

Best Local Similarity 33.0%; Pred. No. 5.5e-06;

Matches 36; Conservative 28; Mismatches 40; Indels 5; Gaps 2;

QY 16 LNVIVDDPLNLIIEKIIKAIGGISTQANNGEAVIIHRDGGSFLLIMDKEMPERD 75

DB 514 LHVLLVEDNATRLTATRMLEALGARVTTAEDGAGQVAAARQG---FDLILMDIQMPVMD 570

QY 76 GVSTTKKLEMEVKSMIVGVTSLADN--EERRAFMEAGLNHCLAKPLT 122

DB 571 GVEATHIRAFNSPAGAAPILAMTANAMAHQAQSYLAAGMDGAIAKPLS 619

RESULT 12

AD0432

aerobic respiration control sensor/response regulatory protein (EC 2.7.3.-) [imported]

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD0432

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0432

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92784.1; PID:g15981477; GSPDB:GN00175

A:Gene: arcB

C:Genetics:

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator ho

C:Keywords: phosphotransferase

Query Match 24.0%; Score 162.5; DB 2; Length 778;

Best Local Similarity 27.0%; Pred. No. 6.8e-06;

Matches 33; Conservative 38; Mismatches 48; Indels 3; Gaps 2;

QY 16 LNVIVDDPLNLIIEKIIKAIGGISTQANNGEAVIIHRDGGSFLLIMDKEMPERD 75

DB 525 LHVLLVEDIELNVIVARSVLEKLGNSVDVANNGHDALAMFNP--EEDFLVLLDIQLPDMS 582

QY 76 GVSTTKKLEMEVKSMIVGVTSLADNE--EERRAFMEAGLNHCLAKPLTKDKIPLINQLM 134

DB 583 GLDIARQIRAEYKGSQSLPPLVALTANVLKDKKEYLDAGMDVLSKPLSVFALTAMIKQFW 642

QY 135 DA 136

DB 643 DS 644

RESULT 13

A69487

response regulator homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Mar-2003  
C:Accession: A69487  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L. Nature 390, 364-370, 1997  
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: A69487  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-117 <KLS>  
A:Cross-references: GB:AE000971; GB:AE000782; NID:g2689294; PIDN:AA89351.1; PID:g264864  
C:Superfamily: signal transduction receiver (phosphoacceptor) protein (CheY-like); resp  
C:Keywords: phosphoprotein  
F:4-110/Domain: response regulator homolog <RR>  
F:50/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 23.5%; Score 159; DB 2; Length 117;  
Best Local Similarity 28.3%; Pred. No. 1.5e-06;  
Matches 36; Conservative 42; Mismatches 31; Indels 18; Gaps 5;

QY 16 LNVLVDD-----DPLNLIHE-KIIRAIGGISQTANNGEAVIIHRDGGSSFDLILMDK 69

DB 2 IKVMVDDIEAMREILKIMDKYVIE-----ASNGREAVELYRE--ERPDVILMDV 51

QY 70 EMPERDGVSTTKKLEMEVKSVMIVGVTSLADNEERAFMEAGLNHCIAKPLTKDKIPL 129

DB 52 MPMPLANGIEATSEIKKIDPAKIVAIYAYASSKGEK--VIEAGADYILKPFTRKEVVEL 109

QY 130 INQLMDA 136

DB 110 IKRIILNS 116

RESULT 14

C82151 sensor histidine kinase Vcl1831 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: C82151

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 405, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: C82151

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-736 <HEI>

A:Cross-references: GB:AE004259; GB:AE003852; NID:g9656353; PIDN:AAF94979.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El for

C:Genetics:

A:Gene: Vcl1831

A:Map position: 1

Query Match 23.5%; Score 159; DB 2; Length 736;  
Best Local Similarity 30.8%; Pred. No. 1.3e-05;  
Matches 41; Conservative 35; Mismatches 45; Indels 12; Gaps 5;

QY 8 DIEKIKK---LNVLVDDPLNLIHEKIKAIGGISQTANNGEAVIIHRDGGSSFD 63

DB 590 DVKPTAKTYPLGLRVILVEDNTNIMLEAFMRNKFECHSVMGVOAITALQE--SSFD 647

QY 64 LILMDKEMPERDGVSTTKKLEMEVKSVMIVGVTSLADNEERAFMEAGLNHCIAK 119

DB 648 LVLMNDHMPKDGIOGATRIOLPLPQAKILLFGCT--ADVFKDTRDKMLSAGADDDIAK 705

QY 120 PLTRDKIIPILNQ 132

DB 706 PIAHELDMALEQ 718

RESULT 15

AG1897

two-component hybrid sensor and regulator all0729 [imported] - Nostoc sp. (strain PCC

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: AG1897

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 203-213, 2001

A:Title: Complete Genomic

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG1897

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1645 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB72686.1; PID:gl7130074; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0729

Query Match 23.5%; Score 159; DB 2; Length 1645;

Best Local Similarity 27.7%; Pred. No. 3.1e-05;

Matches 39; Conservative 40; Mismatches 50; Indels 12; Gaps 3;

DNA Res. 8, 203-213, 2001

A:Title: Complete Genomic

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG1897

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1645 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB72686.1; PID:gl7130074; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0729

Search completed: August 13, 2003, 19:12:39

Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2003, 19:08:31 ; Search time 11 Seconds  
(without alignments)  
581.421 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 578

Sequence: 1 MATKSMGDIEKKKLVLI.....LAKPLTKDKIIPINOLMDA 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	178.5	26.3	1197	1 EVGS_ECOLI	P30855 escherichia
2	176.5	28.0	1197	1 EVGS_ECO57	P58402 escherichia
3	173.5	25.6	948	1 RCSC_SALTY	Q56128 salmonella
4	172.5	25.4	948	1 RCSC_SALTY	P58662 salmonella
5	171.5	25.3	949	1 RCSC_ECOLI	P14376 escherichia
6	156.5	23.1	777	1 FRZE_MYXXA	P18769 myxococcus
7	150	22.1	622	1 SKN7_YEAST	P38889 saccharomyc
8	149.5	22.1	778	1 ARCB_ECO57	P58363 escherichia
9	149.5	22.1	778	1 ARCB_ECOLI	P22763 escherichia
10	145	21.4	1220	1 SUNI_YEAST	P39928 saccharomyc
11	143.5	21.2	712	1 SKL1_YEAST	Q07084 saccharomyc
12	142	20.9	907	1 GACS_PSESY	P48027 pseudomonas
13	140.5	20.7	226	1 DCTR_BACSU	P96602 bacillus su
14	140.5	20.7	230	1 DCTR_BACHD	Q9K998 bacillus ha
15	134	19.8	119	1 CHEY_BACSU	P24072 bacillus su
16	134	19.8	1238	1 BVGS_BORPE	P16575 bordetella
17	133.5	19.7	228	1 DCTR_BACME	P39486 bacillus me
18	133.5	19.7	918	1 BARA_ECOLI	P26607 escherichia
19	132.5	19.5	131	1 YSOL_PLEBO	P51586 plectonena
20	132.5	19.5	918	1 BARA_SHIFL	P59342 shigella fl
21	132	19.5	859	1 LUXQ_VIBHA	P54302 vibrio harv
22	131.5	19.4	441	1 ZRAR_SALTI	Q82333 salmonella
23	131.5	19.4	441	1 ZRAR_SALTV	P25852 salmonella
24	130	19.2	1238	1 BVGS_BORPA	P40330 bordetella
25	128.5	19.0	849	1 LUXN_VIBRA	P4301 vibrio harv
26	127.5	18.8	522	1 MCS4_SCHPO	P87323 schizosacch
27	127	18.7	1238	1 BVGS_BORBR	P26762 bordetella
28	126.5	18.7	227	1 CPXR_HAEIN	P44895 haemophilus
29	125.5	18.5	443	1 ZRAR_KLEOX	Q9apd9 klebsiella
30	124	18.3	144	1 CHEY_TREPA	P96126 treponema p
31	124	18.3	539	1 PRRI_SCHPO	O14283 schizosacch
32	122.5	18.1	441	1 ZRAR_ECOLI	P4375 escherichia
33	122	18.0	122	1 ALGR_PSEAE	P26275 pseudomonas

## RESULT 1

EVGS_ECOLI	STANDARD:	PRT:	1197 AA.
ID	P30855; P77644; Q9RF36; Q9RF37;		
AC	01-JUL-1993 (Rel. 26, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Sensor protein evgs precursor (EC 2.7.3.-)		
GN	EVGS OR B2370.		
OS	Escherichia coli.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=562;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12;		
RX	MEDLINE=94171083; PubMed=8125343;		
RA	Utsumi R., Katayama S., Taniguchi M., Horie T., Ikeda M., Igaki S.,		
RA	Nakagawa H., Miwa A., Tanabe H., Noda M.;		
RT	"Newly identified genes involved in the signal transduction of		
RT	Escherichia coli K-12.";		
RL	Gene 140:73-77(1994).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12;		
RX	MEDLINE=93173621; PubMed=1289796;		
RA	Utsumi R., Katayama S., Ikeda M., Nakagawa H., Miwa A.,		
RA	Taniguchi M., Noda M.;		
RT	"Cloning and sequence analysis of the evgAS genes involved in signal		
RT	transduction of Escherichia coli K-12.";		
RL	Nucleic Acids Symp. Ser. 27:149-150(1992).		
RN	[3]		
RP	SEQUENCE FROM N.A. (VARIANTS EVGS1 AND EVGS4).		
RC	STRAIN=K12;		
RX	MEDLINE=2037813; PubMed=10923791;		
RA	Kato A., Ohnishi H., Yamamoto K., Furuta E., Tanabe H., Utsumi R.;		
RT	"Transcription of emrKY is regulated by the EvgA-Evgs two-component		
RT	system in Escherichia coli K-12.";		
RL	Biosci. Biotechnol. Biochem. 64:1203-1209(2000).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12 / MG1655;		
RX	MEDLINE=97426617; PubMed=9278503;		
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,		
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,		
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,		
RA	Mau B., Shao Y.;		
RT	"The complete genome sequence of Escherichia coli K-12.";		
RL	Science 277:1453-1474(1997).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12;		
RX	MEDLINE=97349980; PubMed=9205837;		
RA	Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,		
RA	Itch T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,		
RA	Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,		

Q8x613 escherichia  
P10957 escherichia  
O34534 bacillus su  
Q45994 cauliobacter  
P16244 escherichia  
O44006 alcaligenes  
P49246 xanthomonas  
Q56312 thermotoga  
P51358 porphyra pu  
P43501 pseudomonas  
O06065 escherichia  
O78428 guillardia

## ALIGNMENTS

Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,  
 Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 Yamagata S., Horiuchi T.,  
 "Construction of a contiguous 874-kb sequence of the *Escherichia coli*  
 - K12 genome corresponding to 50.0-68.8 min on the linkage map and  
 analysis of its sequence features.";  
 DNA Res. 4:91-113(1997).  
 [6]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98194702; PubMed=9535079;  
 RA Petraud A.-L., Kimmel B., Weiss V., Gross R.;  
 "Specificity of the BvgAS and EvgAS phosphorelay is mediated by the  
 C-terminal HPT domains of the sensor proteins.";  
 Mol. Microbiol. 27:875-887(1998).  
 CC -!- FUNCTION: Member of the two-component regulatory system evgs/evgA.  
 CC phosphorylates evgA via a four-step phosphorelay in response to  
 environmental signals.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Probable).  
 CC group from a His in the primary transmitter domain, to an Asp in  
 CC the receiver domain and to a His in the secondary transmitter  
 CC domain.  
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
 CC -!- SIMILARITY: Contains 1 HPT domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D14008; BAA03108.1;  
 DR EMBL; AF201840; AAF17563.1;  
 DR EMBL; AF201841; AAF17564.1;  
 DR EMBL; AE000325; AAC75429.1;  
 DR EMBL; D90867; BAA16241.1;  
 DR PIR; G5010; G65010.  
 DR HSSP; P06143; 1DUR.  
 DR Ecogen; EGI1610; evgs.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003661; His\_KinA.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR001789; Response\_reg.  
 DR InterPro; IPR001311; SBP\_glu\_receptor.  
 DR InterPro; IPR001638; SBP\_bac\_3.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00512; HlsKA; 1.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR Pfam; PF00497; SBP\_bac\_3; 1.  
 DR PRINTS; P000344; BCTELSENSOR.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00388; HlsKA; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00062; PBPB; 2.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 DR PROSITE; PS50894; HPT; 1.  
 DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
 KW Sensory transduction; Transferase; Kinase; Phosphorylation;  
 KW Transmembrane; Inner membrane; Signal; Complete proteome.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1197  
 FT DOMAIN 22 325  
 FT TRANSMEM 346 346  
 FT DOMAIN 347 537  
 FT TRANSMEM 538 558  
 FT TRANSMEM 558

FT DOMAIN 559 1197  
 FT DOMAIN 718 938  
 FT DOMAIN 960 1074  
 FT DOMAIN 1098 1197  
 FT MOD\_RES 721 721  
 FT MOD\_RES 1009 1009  
 FT MOD\_RES 1137 1137  
 FT VARIANT 577 577  
 FT VARIANT 701 701  
 FT CONFLICT 152 152  
 FT CONFLICT 242 243  
 FT CONFLICT 275 275  
 FT CONFLICT 420 421  
 FT CONFLICT 739 739  
 FT CONFLICT 758 758  
 FT CONFLICT 761 761  
 FT CONFLICT 877 877  
 FT CONFLICT 1045 1045  
 FT CONFLICT 1074 1074  
 FT CONFLICT 1197 AA; E8E1DE0F797B1278 CRC64;  
 SQ SEQUENCE 1197 AA; 134742 MW; E8E1DE0F797B1278 CRC64;  
 Query Match 26.3%; Score 178.5; DB 1; Length 1197;  
 Best Local Similarity 34.4%; Pred. No. 2.9e-07;  
 Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;  
 Qy 12 IKKLNVLIVDDPLNLIIHEKIIKAIGGISQTANNGEEAVIHRDGGSSFDLILMDKEM 71  
 Db 955 LPEKLSIIADDDPTNLLKQLNLLGYDVBDEATDGVQA--LHKVSMQHYDLLITDVM 1012  
 Qy 72 PERDGVSTFKLEMEVKSNVGVTSADNEERRAFMEAGLNHCLAKPLTKDKTIPLIN 131  
 Db 1013 PNMGGFELTKLRQNSNPIMGLTANA-QANEREKGLSCGMNLCFLPLILDVLTHTLS 1071  
 Qy 132 QL 133  
 Db 1072 QL 1073  
 RESULT 2  
 EVGS\_ECO57 STANDARD; PRT; 1197 AA.  
 ID EVGS\_ECO57 STANDARD; PRT; 1197 AA.  
 AC P58402;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sensor protein evgs precursor (EC 2.7.3.-).  
 GN EVGS OR Z3632 OR ECS3249.  
 OS *Escherichia coli* O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Escherichia*.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RMD 0509952;  
 RX MEDLINE=2115231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RT



CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; AL627274; CAD07502.1; -	
DR	EMBL; AE016836; AAC68299.1; -	
DR	EMBL; X87830; CRA61095.1; -	
DR	HSSP; P06143; 1UDR.	
DR	InterPro; IPR003594; ATPbind_ATPase.	
DR	InterPro; IPR004358; Bact_sens_pr-C.	
DR	InterPro; IPR003661; His_KinA.	
DR	InterPro; IPR005467; His_Kinase.	
DR	InterPro; IPR000014; PAS_domain.	
DR	InterPro; IPR001789; response_reg.	
DR	Pfam; PF03518; HATPase_C; 1.	
DR	Pfam; PF05512; HiskA; 1.	
DR	Pfam; PF00072; response_reg; 1.	
DR	PRINTS; PR00344; BCTRLSENSOR.	
DR	ProDom; PD000039; Response_reg; 1.	
DR	SMART; SM00387; HATPase_C; 1.	
DR	SMART; SM00388; HiskA; 1.	
DR	SMART; SM00448; REC; 1.	
DR	PROSITE; PS50109; HIS_KIN; 1.	
DR	PROSITE; PS50112; PAS; FALSE_NEG.	
DR	PROSITE; PS50110; RESPONSE_REGULATORY; 1.	
DR	Sensory transduction; transiferase; Kinase; Bacterial capsule;	
KW	Inner membrane; Transmembrane; Phosphorylation; Complete proteome.	
FT	DOMAIN 1 20	
FT	DOMAIN 1 20	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 21 41	POTENTIAL.
FT	DOMAIN 42 313	PERIPLASMIC (POTENTIAL).
FT	TRANSMEM 314 334	POTENTIAL.
FT	DOMAIN 335 948	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 357 425	PAS.
FT	DOMAIN 476 692	HISTIDINE KINASE.
FT	DOMAIN 826 940	RESPONSE REGULATORY.
FT	DOMAIN 826 940	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES 479 479	PHOSPHORYLATION (BY SIMILARITY).
FT	MOD_RES 875 875	
FT	SEQUENCE 948 AA: 106237 MW: AE3A21701265A865 CRC64;	
SO		

Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L.; Miller W., Stoneking T., Nhan M., Waterston K., Wilson R.K.: "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2." Nature 413:852-856(2001).

-! FUNCTION: Member of the two-component regulatory system rcsC/rcsB involved in the regulation of the expression of genes involved in colanic acid capsule synthesis. RcsC probably functions as a membrane-associated protein kinase that phosphorylates rcsB in response to environmental signals.

-! SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Probable).

-! PTM: Activation probably requires a transfer of a phosphate group between a His in the transmitter domain and an Asp of the receiver domain (By similarity).

-! SIMILARITY: Contains 1 histidine kinase domain.

-! SIMILARITY: Contains 1 PAS (PER-ARNY-SIM) dimerization domain.

-! SIMILARITY: Contains 1 response regulatory domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpost at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (<http://www.isb-sib.ch/announcement>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

EMBL; AE008801; AACL21172.1; StyGene; SG7?7??, rcsC InterPro: IPR003594; ATPbind\_ATPase. InterPro: IPR004358; Bact\_sens\_Pt\_C. InterPro: IPR003661; His\_kinA. InterPro: IPR005467; His\_kinase. InterPro: IPR000014; PAS\_domain. InterPro: IPR001789; Response\_reg. Pfam: PF02518; HAIPase\_C\_1. Pfam: PF00312; HiskA\_1. Pfam: PF00072; response\_reg\_1. PRINTS; P000344; CTRLSENSOR. ProDom; PD000039; Response\_reg\_1. SMART; SM00387; HAIPase\_C\_1. SMART; SM00388; Hiska\_1. SMART; SM00448; REC\_1. PROSITE; PS0109; HIS\_KIN\_1. DR PROSITE; PS0112; PAS\_FALSE\_NEG. DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1. KW Sensory transduction; Transferase; Kinase; Bacterial capsule; FT Inner membrane; Transmembrane; Phosphorylation; Complete proteome. FW Domain 1 20 CYTOPLASMIC POTENTIAL). FT TRANSMEM 21 41 POTENTIAL. FT DOMAIN 42 313 PERIPLASMIC POTENTIAL). FT TRANSMEM 314 334 POTENTIAL. FT DOMAIN 335 948 CYTOPLASMIC POTENTIAL). FT DOMAIN 357 425 PAS. FT DOMAIN 476 692 HISTIDINE KINASE.. FT DOMAIN 826 940 RESPONSE REGULATORY. FT MOD\_RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY). FT MOD\_RES 875 875 PHOSPHORYLATION (BY SIMILARITY). SQ SEQUENCE 948 AA; 106279 MW; BAAD8DA557D5868B CRG64;

Query Match 25.4%; Score 172.5; DB 1; Length 948;  
Best Local Similarity 34.6%; Pred. No. 7.le-07;  
Matches Conservative 32; Mismatches 35; Indels 3; Gaps 2

18 VLTVDDPDNLNLIHEKIIKAIGGSQATANNGEAAVTHRDGGSGFLLMDXEMPGRGV 77  
::|||::: ||::|::|::|::|:  
827 LVVDHPNTNRRLADOLGLSYGYCKTKANDCVDALNYL--SKNAIDLVLSDMPNMNDGY 884  
::|||::: ::|||::|::|::|:

QY 78 STTKKUREMEVKSMTIVGTSIADNEEPRRAFMEAGLNHCLAKPLTKD 124  
| : : : : : | : : : : | : : : : |  
Db 885 RLTORIQGLGLPVGVTANALAEQRCL-LESGMDSCLSKPVTLTD 930



01-NOV-1990 (Rel. 16, Last sequence update)  
 28-FEB-2003 (Rel. 41, Last annotation update)  
 Gliding motility regulatory protein (EC 2.7.3.-).  
 FRZE.

OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;  
 OC Cytochromatiales; Myxococcaceae; Myxococcus.  
 OX NCBI\_TaxID=34;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=90332690; PubMed=2165608;  
 RA McCleary W.R., Zusman D.R.;  
 RT "FrzE of Myxococcus xanthus is homologous to both CheA and CheY of  
 RL Salmonella typhimurium";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5898-5902(1990).  
 [2]  
 RN PHOSPHORYLATION OF HIS-49.  
 RP MEDLINE=91072208; PubMed=2123853;  
 RA McCleary W.R., Zusman D.R.;  
 RT "Purification and characterization of the Myxococcus xanthus FrzE  
 RL protein shows that it has autophosphorylation activity";  
 RL J. Bacteriol. 172:6661-6668(1990).  
 CC -!- FUNCTION: FRZE IS INVOLVED IN A SENSORY TRANSDUCTION PATHWAY THAT  
 CC CONTROLS THE FREQUENCY AT WHICH CELLS REVERSE THEIR GLIDING  
 CC DIRECTION. FRZE SEEMS TO BE CAPABLE OF AUTOPHOSPHORYLATING ITSELF  
 CC ON AN HISTIDINE RESIDUE AND THEN TO TRANSFER THAT GROUP TO AN  
 CC ASPARTATE RESIDUE IN THE C-TERMINAL PART OF THE PROTEIN.  
 CC -!- SIMILARITY: Contains 1 cheW-like domain.  
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.  
 CC -!- SIMILARITY: Contains 1 HPT domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; M35192; AAA25396.1;  
 PIR; A35966; A35966.  
 DR HSP: O56310; 1B30.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro: IPR002545; Chew.  
 DR InterPro: IPR005467; His\_kinase.  
 DR InterPro: IPR002570; Hpt.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF01584; Chew; 1.  
 DR Pfam: PF02518; HATPase\_c; 1.  
 DR Pfam: PF01627; Hpt; 1.  
 DR Pfam: PF00072; response\_reg; 1.  
 DR PRINTS: PR00344; BCTRLSENSOR.  
 DR ProDom: PD000039; Response\_reg; 1.  
 DR SMART: SM00260; Chew; 1.  
 DR SMART: SM00387; HATPase\_c; 1.  
 DR SMART: SM00073; HPT; 1.  
 DR SMART: SM00448; REC; 1.  
 DR PROSITE: PS00851; CHEW; 1.  
 DR PROSITE: PS0109; HIS\_KIN; 1.  
 DR PROSITE: PS00894; HPT; 1.  
 DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.  
 KW Sensory transduction; Transferase; Kinase; Phosphorylation.  
 FT DOMAIN 1 108 HPT  
 FT DOMAIN 270 509 HISTIDINE KINASE.  
 FT DOMAIN 511 645 CHEW-LIKE.  
 FT DOMAIN 660 776 RESPONSE REGULATORY.  
 FT MOD\_RES 49 49 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT DOMAIN 130 197 ALA/PRO-RICH (POSSIBLE HINGE REGION).  
 CC SEQUENCE 777 AA; 83189 MW; 9912BD40991C69E5 CRC64;

Query Match

23.1%; Score 156.5; DB 1; Length 777;

Best Local Similarity 35.2%; Pred. No. 1.2e-05;  
 Matches 38; Conservative 27; Mismatches 38; Indels 5; Gaps 3;  
 QY 14 KKLNVLIIVDDDDPLLIITHEKIIKAIGISQTANGGEAVIIRHGGSPDLILMKEMPE 73  
 DB 657 KRLRVLLVDDSPITARATEGALVKALGHSVEAQDGEAYV--KVQNTYDLILTDVQMPK 714  
 QY 74 RDGVSTTKKLREMEVKSNI--GVVTSIADNEERRAFMEAGLNHCLAK 119  
 DB 715 LDGFSLARLUKSTPAVARIPVILSLASPEDKRRG--LDAGADAYLVK 761  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=94042854; PubMed=82266533;  
 RA Brown J.L., North S., Bussey H.;  
 RT "SKN7, a yeast multicopy suppressor of a mutation affecting cell wall  
 RT beta-glucan assembly, encodes a product with domains homologous to  
 RT prokaryotic two-component regulators and to heat shock transcription  
 RT factors";  
 RT J. Bacteriol. 175:6908-6915(1993).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP Krens B., Charizanis C., Entian K.-D.;  
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston V.,  
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 RA Nham M., Rifkin L., Riles L., St. Peter H., Trevisan E., Vaughan K.,  
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaudin M.;  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII";  
 RL Science 265:2077-2082(1994).  
 RN [4]  
 RN FUNCTION, AND MUTAGENESIS.  
 RX MEDLINE=95045411; PubMed=7957083;  
 RA Brown J.L., Bussey H., Stewart R.C.;  
 RT "Yeast Skn7p functions in a eukaryotic two-component regulatory  
 RT pathway";  
 RL EMBO J. 13:5186-5194(1994).  
 CC -!- FUNCTION: INVOLVED IN OXIDATIVE STRESS. TRANSCRIPTION FACTOR THAT  
 CC MAY FUNCTION IN A TWO-COMPONENT SIGNAL TRANSDUCTION PATHWAY THAT  
 CC ACTS IN PARALLEL WITH THE PKC1 CASCADE TO REGULATE GROWTH AT THE  
 CC CELL SURFACE.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

-----
DR EMBL: U00485; AAC48911.1; -
DR EMBL: X83031; CAAS8143.1; -
DR EMBL: U00029; AAB69734.1; -
DR PIR: A49344; A49344.
DR HSP: P22121; 2HTS.
DR TRANSFAC: T03481; -.
DR SGD: S0001249; SKN7.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0003700; F:transcription factor activity; IDA.
DR GO: GO:000156; F:two-component response regulator activity; IDA.
DR GO: GO:0006350; P:transcription; IDA.
DR InterPro: IPR000232; HSF_DNA_BIND.
DR InterPro: IPR002341; HSF_ETs.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00447; HSF_DNA_bind; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00056; HSFDOMAIN.
DR ProDom: PD001788; HSF_DNA_bind; 1.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00415; HSF; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS00434; HSF DOMAIN; 1.
DR PROSITE: PS01110; RESPONSE_REGULATORY; 1.
KW Transcription regulation; Sensory transduction; Nuclear protein;
KW DNA-binding; Phosphorylation.
FT DNA_BIND 86 190 BY SIMILARITY.
FT DOMAIN 378 492 RESPONSE REGULATORY.
FT MOD_RES 427 427 PHOSPHORYLATION (PROBABLE).
FT MUTAGEN 427 427 D->N: DIMINISHED ACTIVITY.
FT MUTAGEN 427 427 D->E: AUGMENTED ACTIVITY.
SQ SEQUENCE 622 AA; 69202 MW; 4C732PD66B326742 CRC64;
Query Match 22.1%; Score 150; DB 1; Length 622;
Best Local Similarity 28.1%; Pred. No. 3.3e-05;
Matches 38; Conservative 32; Mismatches 45; Indels 20; Gaps 4;
QY 12 IKKLNVLVDLDDPLNLIHEIKIKAIGGISQTANNGEAVII---HRDGSSFDLIIMD 68
* Db 373 LRKGFHVLVEDDANSVQLCSFLRYGCTGVVSDGLSAISTLEKYR-----YDLVIMD 427
QY 69 KEMPERDGVSTTKKLEMEVKSMVGVVSLANDEERRAFMEAGNLHCLAPLTKDKI-- 126
Db 428 IVPNLEGATATSVRSFONETPIIATGNTGNIMN-ODLITYLQHGMDILARPFTRDLHS 486
QY 127 -----IPLINQ 132
Db 487 ILIRYLKDRIPICEQ 501
RESULT 8
ID ARCB_ECO57 STANDARD; PRT; 778 AA.
AC P88363;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aerobic respiration control sensor protein arcb (EC 2.7.3.-).
GN ARCB OR 24574 OR ECS4089.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posral G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

```

```

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Murata T., Tanaka M., Tobe T.,
RA Han C.-G., Ohtsubo E., Nakayama K., Ogasawara N., Yasunaga T.,
RA Iida T., Takami H., Honda T., Sasekawa C., Shinagawa H.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Member of the two-component regulatory system arcb/arca.
CC Sensor-regulator protein for anaerobic repression of the arc
CC modulon. Activates arca via a four-step phosphorelay. ArcB can
CC also dephosphorylate arca by a reverse phosphorelay involving His-
CC 717 and Asp-576 (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 HPT domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: AEO05549; AAG58344.1; -.
DR EMBL: AP002564; BAB37512.1; -.
DR PIR: A91140; A91140.
DR PIR: D85985; D85985.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR002570; Hpt.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; HSKA; 1.
DR Pfam: PF00989; PAS; 1.
DR Pfam: PF00072; response_reg; 1.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HSKA; 1.
DR SMART: SM00073; HPT; 1.
DR SMART: SM00091; PAS; 1.
DR SMART: SM00448; REC; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0894; HPT; 1.
DR PROSITE: PS0894; HPT; 1.
DR PROSITE: PS0113; PAC; 1.
DR PROSITE: PS0112; PAS; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Transference; Kinase; Phosphorylation;
KW Transmembrane; Inner membrane; Transcription regulation;
KW Complete proteome.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 26 46 POTENTIAL.
FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 153 223 PAS.
FT DOMAIN 226 278 PAC.
FT DOMAIN 289 507 HISTIDINE KINASE.
FT DOMAIN 527 643 RESPONSE REGULATORY.
FT DOMAIN 678 771 HPT.
FT MOD_RES 292 292 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 576 576 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 717 717 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 778 AA; 88010 MW; C8AE004B007F9D30 CRC64;

Query Match 22.1%; Score 149.5; DB 1; Length 778;
Best Local Similarity 25.6%; Pred. No. 4 6e-05;
Matches 31; Conservative 40; Mismatches 47; Indels 3; Gaps 2;

QY 16 LNVLIYDDPLNLIIHKIKIKAIGGISQTANNGEAVIHRDGGSSFDILDMKEMPERD 75
DB 526 LNVLIYDDPLNLIIHKIKIKAIGGISQTANNGEAVIHRDGGSSFDILDMKEMPERD 583
QY 76 GVSUTKKIRENEKSMVTVGVTSLADNB-EERRAFMEAGLNHCIAKPLTKDKIIPLNOLM 134
DB 584 GLDLSRELTKKYPRDPLVLTANVLKQEVNLNAGDDVLSKPLSPALTAIKKFW 643
QY 135 D 135
DB 644 D 644

RESULT 9
ARCH_ECOLI STANDARD; PRT; 778 AA.
AC P22763;
DT 01-AUG-1991 (Rel. 19, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aerobic respiration control sensor protein arCB (EC 2.7.3.-).
GN ARCB OR B3210 OR SF3250.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=90355832; PubMed=2201868;
RA Iuchi S., Matsuda Z., Fujiwara T., Lin E.C.C.;
RT "The arCB gene of Escherichia coli encodes a sensor-regulator protein
RL for anaerobic repression of the arc regulon."
RN [2] Mol. Microbiol. 4:715-727(1990).
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RN [3] Science 277:1232-1244(1997).
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Ou D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RN through comparison with genomes of Escherichia coli K12 and O157."
RN [4] Nucleic Acids Res. 30:4432-4441(2002).
RP CHARACTERIZATION.

```

```

RC SPECIES=E.coli; STRAIN=MJ5;
RX MEDLINE=37431492; PubMed=9286997;
RA Georgellis D., Lynch A.S., Lin E.C.C.;
RT "In vitro phosphorylation study of the arc two-component signal
RL transduction system of Escherichia coli."
RN [5] J. Bacteriol. 179:5429-5435(1997).
RP CHARACTERIZATION.
RC SPECIES=E.coli; STRAIN=MJ5;
RX MEDLINE=99047671; PubMed=9830034;
RA Georgellis D., Kwon O., De Wulf P., Lin E.C.C.;
RT "Signal decay through a reverse phosphorelay in the arc two-component
RL signal transduction system."
RN [6] J. Biol. Chem. 273:32864-32869(1998).
RC MUTAGENESIS OF HIS-292; ASP-576 AND HIS-717.
RX SPECIES=E.coli; STRAIN=K12 / MC4100;
RX MEDLINE=20309722; PubMed=10851007;
RA Kwon O., Georgellis D., Lin E.C.C.;
RT "Phosphorelay as the sole physiological route of signal transmission
RL by the arc two-component system of Escherichia coli."
RN [7] J. Bacteriol. 182:3858-3862(2000).
RC X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS) OF 660-778.
RX SPECIES=E.coli;
RX MEDLINE=97207018; PubMed=9054511;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Insights into multistep phosphorelay from the crystal structure of
RL the C-terminal HPT domain of ArcB."
RN [8] Cell 88:717-723(1997).
RC X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 659-776 IN COMPLEX WITH CHEY.
RX SPECIES=E.coli;
RX MEDLINE=98437504; PubMed=9761838;
RA Kato M., Mizuno T., Hakoshima T.;
RT "Crystallization of a complex between a novel C-terminal transmitter,
RL HPT domain, of the anaerobic sensor Kinase ArcB and the chemotaxis
RN [9] response regulator CheY."
RX Acta Crystallogr. D 54:140-142(1998).
RC X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 659-776.
RX SPECIES=E.coli;
RX MEDLINE=2003135; PubMed=10531481;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Refined structure of the histidine-containing-phosphotransfer (HPT)
RL domain of the anaerobic sensor Kinase ArcB from Escherichia coli at
RN 1.57-A resolution."
RX Acta Crystallogr. D 55:1842-1849(1999).
CC -!- FUNCTION: Member of the two-component regulatory system arcB/arCB.
CC Sensor-regulator protein for anaerobic repression of the arc
CC modulon. Activates arcA via a four-step phosphorelay. ArcB can
CC also dephosphorylate arcA by a reverse phosphorelay involving His-
CC 717 and Asp-576.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 HPT domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

DR EMBL: X53315; CAA37397.1; -  
DR EMBL: U18997; AAB58012.1; -  
DR EMBL: AE000400; AAC76242.1; -  
DR EMBL: AE015336; AAN44715.1; -  
DR PIR: D65112; RGEAR.  
DR PDB: 1A0B; 18-MAR-98.  
DR PDB: 2A0B; 17-JUN-98.  
DR PDB: 1BDJ; 11-MAY-99.  
DR PDB: 1FR0; 31-DEC-02.  
DR EcoGen: EG10062; atcB.  
DR InterPro: IPR003594; Atbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003661; His\_kinA.  
DR InterPro: IPR005467; His\_kinase.  
DR InterPro: IPR002570; Hpt.  
DR InterPro: IPR000700; PAS-assoc\_C.  
DR InterPro: IPR000014; PAS\_domain.  
DR InterPro: IPR001789; Response\_reg.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; HsK; 1.  
DR Pfam: PF00989; PAS; 1.  
DR Pfam: PF00072; response\_reg; 1.  
DR PRINTS; PD00344; BCTLSNSOR.  
DR ProDom: PD000039; Response\_reg; 1.  
DR SMART; SM00388; HsK; 1.  
DR SMART; SM00073; HPT; 1.  
DR SMART; SM00091; PAS; 1.  
DR SMART; SM00448; REC; 1.  
DR TIGRFS; TIGR00229; sensory\_box; 1.  
DR PROSITE; PS0109; HIS\_KIN; 1.  
DR PROSITE; PS0894; HPT; 1.  
DR PROSITE; PS0113; PAC; 1.  
DR PROSITE; PS0112; PAS; 1.  
DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
DR Sensory transduction; Transducer; Kinase; Phosphorylation;  
DR Transmembrane; Inner membrane; Transcription regulation; 3D-structure;  
DR Complete proteome.  
DR DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).  
DR TRANSFEM 26 46 POTENTIAL.  
DR DOMAIN 47 57 PERIPLASMIC (POTENTIAL).  
DR TRANSFEM 58 78 POTENTIAL.  
DR DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).  
DR DOMAIN 153 223 PAS.  
DR DOMAIN 226 278 PAC.  
DR DOMAIN 289 507 HISTIDINE KINASE.  
DR DOMAIN 527 643 RESPONSE REGULATORY.  
DR DOMAIN 678 771 HPT.  
DR MOD\_RES 292 292 PHOSPHORYLATION (AUTO-).  
DR MOD\_RES 576 576 PHOSPHORYLATION (PROBABLE).  
DR MUTAGEN 292 292 H->Q: LOSS OF ACTIVITY.  
DR MUTAGEN 576 576 D->A: LOSS OF ACTIVITY.  
DR MUTAGEN 717 717 H->Q: LOSS OF ACTIVITY.  
DR CONFLICT 469 470 MISSING (IN REF. 2).  
DR SEQUENCE 778 AA; 87982 MW; DD61EA6ECF95AD30 CRC64;  
Query Match 22.1%; Score 149.5; DB 1; Length 778;  
Best Local Similarity 25.6%; Pred. No. 4.6e-05;  
Matches 31; Conservative 40; Mismatches 47; Indels 3; Gaps 2;  
QY 16 LNVLLVDDPLNLITHEKTIKAGIGISOTANNGEAVIHRDGGSPDLIMDKEMPERD 75  
DB 526 LNVLLVEDLELVNIVARSVLEKLSGSDVMTGKAALENFPG--EYDLVLLDQLPDMT 583  
QY 76 GVSTTKKLREMEYKMGVGTSLADNE-EERRAFMEAGLNHCLAKPLTKDKIIFLINOLM 134  
DB 584 GLDISRELTKRYPRDPLPLVALTANVLKDKQEYLNAGMDVLSPLSVPAITAMIKKFW 643  
QY 135 D 135  
DB 644 D 644

RESULT 10  
SLN1\_YEAST  
ID SLN1\_YEAST STANDARD; PRT; 1220 AA.  
AC P3928;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Osmolarity two-component system protein SLN1 (EC 2.7.3.-).  
GN SLN1 OR YPD2 OR YIL147C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / YPH1;  
RX MEDLINE=94024010; PubMed=8211183;  
RA Ota I.M., Varshavsky A.;  
RT "A yeast protein similar to bacterial two-component regulators.";  
RL Science 262:566-569(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX PubMed=9169870;  
RA Churche C.M., Bowman S., Badcock K., Bankier A., Brown D.,  
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,  
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,  
RA Moulton S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,  
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";  
RL Nature 387:84-87(1997).  
RN [3]  
RP MUTAGENESIS OF HIS-576 AND ASP-1144.  
RX MEDLINE=94239498; PubMed=8183345;  
RA Maeda T., Margler-Murphy S.M., Saito H.;  
RT "A two-component system that regulates an osmosensing MAP kinase  
cascade in yeast.";  
RL Nature 369:242-245(1994).  
CC -1- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM  
CC SLN1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE  
CC EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->  
CC PBS2->HOG1 PATHWAY. INACTIVE SLN1 ALLOWS THE UNPHOSPHORYLATED SSK1  
CC PROTEIN TO ACTIVATE SSK2 AND SSK22, TWO MARKERS THAT FURTHER  
CC STIMULATE THE PBS2-HOG1 MAPK CASCADE. IN LOW OSMOLARITY MEDIA,  
CC THE ACTIVATED SLN1 HISTIDINE KINASE REPRESSSES THE ACTIVATION OF THE  
CC PBS2-HOG1 KINASE CASCADE THROUGH PHOSPHORYLATION OF SSK1.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- PTM: ACTIVATION PROBABLY REQUIRES A TRANSFER OF A PHOSPHATE GROUP  
CC BETWEEN A HIS IN THE TRANSMITTER DOMAIN AND AN ASP OF THE RECEIVER  
CC DOMAIN.  
CC -1- SIMILARITY: Contains 1 histidine kinase domain.  
CC -1- SIMILARITY: Contains 1 response regulatory domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: Z38059; CAA86131.1; -  
CC EMBL: U01835; AAC48912.1; -  
CC PIR: S48387; S48387.  
CC HSP: S001409; SLN1.  
CC SGD: S0001409; SLN1.  
CC GO: GO:0007234; P: osmosensory signaling pathway via two-compo. . . ; IDA.  
CC GO: GO:0042542; P: response to hydrogen peroxide; IMP.  
CC InterPro: IPR003594; Atbind\_ATPase.  
CC InterPro: IPR004358; Bact\_sens\_pr\_C.  
CC InterPro: IPR003661; His\_kinA.  
CC InterPro: IPR005467; His\_kinase.  
CC InterPro: IPR001789; Response\_reg.

```
DR PFam; PF02518; HATPase_c; 1.
DR PFam; PF00512; HSKA; 1.
DR PFam; PF00072; response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PRODom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01109; HIS_KIN; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Transferase; Kinase; Phosphorylation;
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 46 POTENTIAL.
FT DOMAIN 47 333 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 334 354 POTENTIAL.
FT DOMAIN 335 1220 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 573 928 HISTIDINE KINASE.
FT DOMAIN 1089 1210 RESPONSE REGULATION.
FT MOD_RES 576 576 PHOSPHORYLATION (AUTO-) (PROBABLE).
FT MOD_RES 1144 1144 PHOSPHORYLATION (PROBABLE).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 576 576 H->O: INACTIVE.
FT MUTAGEN 891 891 G->D: SLOW GROWTH, SLN1-1 MUTANT.
FT MUTAGEN 1144 1144 D->N: INACTIVE.
SQ SEQUENCE 1220 AA; 134434 MW; 45FFE24A8165486B CRC64;

Query Match 21.4%; Score 145; DB 1; Length 1220;
Best Local Similarity 29.0%; Pred. No. 0.00019;
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DIEKTKKLVNLIYDDPLNLIHEKLIKA---IGGIS--QTANNGEA---VHIRDGG 59
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1080 DDKNETSVKILYVEDNHVN---OEVIKRLNLEGIENTELACDQGEADFKYKELTSKG 1135
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 SFFDILMDKEMPERDGVSTTKKL-REMEVKSMIVGVTSIADNEEERAFMEAGLNHCLA 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1136 ENYNIEMFDMQPKYDGLSTKMIRDLGYTSPIVALTAADDSNIKEC-LESSMNGELS 1194
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 KPLTKDKLIPLINQMDA 136
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1195 KPIKRPKLTILTEPCAA 1212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
SSK1_YEAST STANDARD; PRT; 712 AA.
AC Q07084; Q07909;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Osmolarity two-component system protein SSK1.
GN SSK1 OR YLR006C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=94239498; PubMed=8183345;
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RL cascade in yeast.";
RL Nature 369:242-245(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
```

28-FEB-2003 (Rel. 41, Last annotation update)  
 DN Sensor protein gacS (EC 2.7.3.-).  
 GN GACS OR LENA.  
 OS Pseudomonas syringae (pv. syringae).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 CC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=321;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RN MEDLINE=92234961; PubMed=1314807;  
 RX Arabak E.M., Willis D.K.;  
 RT "The lna gene required for pathogenicity of Pseudomonas syringae pv.  
 RT syringae on bean is a member of a family of two-component  
 RT regulators.";  
 RL J. Bacteriol. 174:3011-3020(1992).  
 CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM  
 CC GACA/GACS(LENA). MAY BE INVOLVED IN LESION FORMATION, SWARMING AND  
 CC IN THE PRODUCTION OF EXTRACELLULAR PROTEASE, SYRINGOMYCIN AND N-  
 CC ACYL-L-HOMOSERINE LACTONE (ACYL-HSL). REQUIRED FOR PATHOGENICITY  
 CC ON BEAN.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Probable).  
 CC -!- PTM: Activation requires a sequential transfer of a phosphate  
 CC group from a His in the primary transmitter domain, to an Asp in  
 CC the receiver domain and to a His in the secondary transmitter  
 CC domain (By similarity).  
 CC -!- SIMILARITY: Contains 1 HAMP domain.  
 CC -!- SIMILARITY: Contains 1 Histidine Kinase domain.  
 CC -!- SIMILARITY: Contains 1 HPT domain.  
 CC -!- SIMILARITY: Contains 1 response regulatory domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation in  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M80477; AAA25877.1; -.  
 DR HSSP; P06143; 1UDR.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003860; HAMP.  
 DR InterPro; IPR003661; His\_kinA.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF00672; HAMP; 1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00312; HiskA; 1.  
 DR Pfam; PF01827; Hpt; 1.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR PRINTS; PR00344; BCTRLSENSOR.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00304; HAMP; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00388; HiskA; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS00865; HAMP; 1.  
 DR PROSITE; PS01009; HIS\_KIN; 1.  
 DR PROSITE; PS00894; HPT; 1.  
 DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
 KW Sensory transduction; Transferase; Kinase; Phosphorylation;  
 KW Transmembrane; Inner membrane.  
 FT TRANSMEM 9 25 POTENTIAL.  
 FT TRANSMEM 84 101 POTENTIAL.  
 FT TRANSMEM 159 178 POTENTIAL.  
 FT DOMAIN 182 234 HAMP.  
 FT DOMAIN 281 502 HISTIDINE KINASE.  
 FT DOMAIN 658 777 RESPONSE REGULATORY.  
 FT DOMAIN 814 907 HPT.  
 FT FT





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2003, 19:09:22 ; Search time 37 Seconds  
(without alignments)  
948.518 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678  
Sequence: 1 MATKSMGDIKIKKLNVL.....LAKPLTKDKIIPLINQMDA 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	678	100.0	136	10	O82445
2	538	79.4	142	10	Q8M84
3	181	26.7	1969	5	O15763
4	179.5	26.5	820	16	Q8Z63
5	178.5	26.3	1197	16	Q8CV5
6	174	25.7	394	16	Q9A30
7	173.5	25.6	949	16	Q8FF9
8	173	25.5	417	2	Q9RLC7
9	171.5	25.3	933	16	Q8XE39
10	171.5	25.3	957	16	Q8ZGR4
11	171	25.2	927	2	Q9ANY0
12	170	25.1	769	16	Q8PQ37
13	167.5	24.7	1364	16	Q8PJN8
14	166.5	24.6	1507	2	O85663
15	163.5	24.1	1364	16	Q8P883
16	162.5	24.0	642	16	Q9A3L2

## ALIGNMENTS

### RESULT 1

O82445 PRELIMINARY: PRT: 136 AA.  
AC O82445;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Response regulator protein.  
OS Brassica napus (rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Whitelaw C.A., Paul W., Jenkins E.S., Taylor V.M., Roberts J.A.;  
RT "A mRNA encoding a response regulator protein from Brassica napus is up-regulated during pod development.";  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF057027; AAC62225.1;  
DR InterPro; IPR001789; Response\_reg.  
DR Pfam; PF00072; Response\_reg; 1.  
DR ProDom; PD000039; Response\_reg; 1.  
DR SMART; SM00448; REC; 1.  
DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
SQ Phosphorylation; Sensory transduction.  
KW SEQUENCE 136 AA; 15035 MW; DA99B768FAB3CFF5 CRC64;

Query Match 100.0%; Score 678; DB 10; Length 136;

Best Local Similarity 100.0%; Pred. No. 4.7e-52;

Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKSMGDIKIKKLNVLIVDDPLNLIIEHKIKAIGGISQTANNGEAAVTHRDGGS 60

|||||

Db 1 MATKSMGDIKIKKLNVLIVDDPLNLIIEHKIKAIGGISQTANNGEAAVTHRDGGS 60

|||||

QY 61 SFDLILMDKEMPERGVSTTKKLREWEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKP 120

|||||

Db 61 SFDLILMDKEMPERGVSTTKKLREWEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKP 120

|||||

```

QY 121 LTKDKIIPILNQLMDA 136
DB 121 LTKDKIIPILNQLMDA 136

RESULT 2
Q9M8Y4
ID Q9M8Y4 PRELIMINARY; PRT; 142 AA.
AC Q9M8Y4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative response regulator protein (receiver component).
DE T6K12.10.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Rowing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC T6K12 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RA "Full-length messenger RNA sequences greatly improve genome
RA annotation.";
RL Genome Biol. 0:0-0(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RA "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC016829; AAF26786.1;
DR EMBL: AY085638; AAM62859.1;
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00072; Response_reg.1.
DR ProDom: PD00039; Response_reg.1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Phosphorylation; Sensory transduction.
SQ SEQUENCE 142 AA; 15732 MW; 57487135FF2D5238 CRC64;

Query Match 79.4%; Score 538; DB 10; Length 142;
Best Local Similarity 78.9%; Pred. No. 9.8e-40;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 1 MATKSGDIEK-----IKKKL-NVLIVDDPLNLIIEHKIKAIGGISQTANNGEEAVII 54
DB 1 MATKSGTETKSTIEVKKKLINVLIVDDPLNRLHEMIKIITGISTQAKNGEEAVIL 60

QY 55 HRDGGSSFDLIIMDKEMPERGVSTTKKLREMEKSMIVGVTSIADNEERRAFMEAGLN 114
DB 61 HRDGEASFDLIIMDKEMPERGVSTTKKLREMKVTSIMVGVTSIADQEEERKAFMEAGLN 120

QY 115 HCLAKPLTKDKIIPILNQLMDA 136
DB 121 HCLAKPLTKAKIFPLISHLFLDA 142

RESULT 3
O15763
ID O15763 PRELIMINARY; PRT; 1969 AA.
AC O15763;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hybrid histidine kinase DHKB.
GN DHKB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAX3;
RX MEDLINE=98248997; PubMed=9576830;
RA Zinda M.J., Singleton C.K.;
RT "The hybrid histidine kinase dhkb regulates spore germination in
RT Dictyostelium discoideum.";
RL Dev. Biol. 196:171-183(1998).
DR EMBL: AF024654; AAB1889.1;
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR Pfam: PF00072; response_reg.1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg.1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; Hiska; 1.
DR SMART: SM00091; PAS; 1.
DR SMART: SM00448; REC; 1.
DR TIGRFAMS: TIGR00229; sensory_box; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 1969 AA; 219025 MW; 8E7A7952A31BB52B CRC64;

Query Match 26.7%; Score 181; DB 5; Length 1969;
Best Local Similarity 35.8%; Pred. No. 5.1e-07;
Matches 48; Conservative 29; Mismatches 39; Indels 18; Gaps 5;

QY 10 EKIKKKLVIVDDPLNLIIEHKIKAIGGISQTANNGEEAVIIHRDGGSSFDLIIMDK 69
DB 1836 EKIEK---ILLVEDFNVNKIFSKLLDKSGYIFDVAHNGVEAVCYKKG--AYDLILMDC 1990

QY 70 EMPERDGVSTTKKLREMEKSMI-----VGVTSLADNE--ERRAFMEAGLNHCL 117
DB 1891 OMPMDGFEATTAIRELEKSNLIESPPSKKSHVIVALTANSYKDKQKCLSVGMNDFL 1950

QY 118 AKPL-TKDKIIPILI 130
DB 1951 QKPIKTSILIQMI 1964

RESULT 4
Q8EZ63
ID Q8EZ63 PRELIMINARY; PRT; 820 AA.
AC Q8EZ63;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN LA3996.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE011554; AAN51193.1;
KW Complete proteome.
SQ SEQUENCE 820 AA; 94325 MW; 3DBE99966794E67D CRC64;

```



RT of uropathogenic *Escherichia coli*.  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
 DR EMBL: AF016763; AAN81215.1; -.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 945 AA; 106590 MW; DF8CA47F9EBE4088 CRC64;

Query Match 25.6%; Score 173.5; DB 16; Length 949;  
 Best Local Similarity 33.6%; Pred. No. 9.7e-07;  
 Matches 37; Conservative 35; Mismatches 33; Indels 5; Gaps 3;  
 QY 18 VLIVDDPDLNLIHKKIATGGISQTANNGEEAV-IHRDGGSSFDLILMDKEMPERDG 76  
 Db 827 ILVVDHPINRLLADQGLSGYCKTANDGVDALNVLK---NHIDIVLSDVNNPNDMG 883  
 QY 77 VSTTKKLREMEVKSMTGVTSIADNEERRAFMEAGLNHCLAKPLTKDKI 126  
 Db 884 YRLTQRIQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDDVI 932

RESULT 8  
 Q9RLC7 PRELIMINARY; PRT; 417 AA.  
 AC Q9RLC7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Putative histidine kinase (Fragment).  
 GN GACS.  
 OS *Pseudomonas stutzeri* (*Pseudomonas perfectomarina*).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 .OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=316;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JM300;  
 RA Graupner S., Wackernagel W.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ249741; CAB56474.1; -.  
 DR InterPro: IPR002570; Hpt.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF00072; response\_reg.1.  
 DR ProDom: PD000039; Response\_reg.1.  
 DR SMART: SM00073; HPT; 1.  
 DR SMART: SM00448; REC; 1.  
 DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.  
 DR Kinase; Phosphorylation; Sensory transduction.  
 FT NON\_TER  
 SQ SEQUENCE 417 AA; 45749 MW; D196B4FFDDA23BB2 CRC64;

Query Match 25.5%; Score 173; DB 2; Length 417;  
 Best Local Similarity 32.8%; Pred. No. 4e-07;  
 Matches 39; Conservative 35; Mismatches 37; Indels 8; Gaps 3;  
 QY 17 NVLIVDDPDLNLIHKKIATGGISQTANNGEEAV-IHRDGGSSFDLILMDKEMPERDG 76  
 Db 169 SVLVCDDNPANLMLLETLITDGGVAVSSQQALEVQQ-QSFDNVFVQVMPGMDG 226  
 QY 77 VSTTKKLREMEVKS----MIVGVTSIADNEERRAFMEAGLNHCLAKPLTKDKI 130  
 Db 227 ROTTEAIRWELESQGPLPVALTAHA-LSNERSLQSLGLDYLTKPTISEROLAQV 284

RESULT 9  
 Q8XE39 PRELIMINARY; PRT; 933 AA.  
 AC Q8XE39;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Sensor for ctr capsule biosynthesis, probable histidine kinase  
 DE acting on RcsB.  
 GN RSCC OR Z3477 OR RCS3107.  
 OS *Escherichia coli* O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Escherichia*.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7."  
 RL Nature 409:529-533 (2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22 (2001).  
 DR EMBL: AF005454; AAC57353.1; -.  
 DR EMBL: AF002560; BAB36330.1; -.  
 DR HSP; P08143; 1b42.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro: IPR003661; His\_kinase.  
 DR InterPro: IPR005467; His\_kinase.  
 DR InterPro: IPR000014; PAS\_domain.  
 DR InterPro: IPR001789; Response\_reg.  
 DR Pfam: PF0518; HATPase\_c; 1.  
 DR Pfam: PF00512; Hiska; 1.  
 DR Pfam: PF00072; response\_reg.1.  
 DR PRINTS: PR00344; BCTRLSENSOR.  
 DR ProDom: PD000039; Response\_reg.1.  
 DR SMART: SM00387; HATPase\_c; 1.  
 DR SMART: SM00388; Hiska; 1.  
 DR SMART: SM00091; PAS; 1.  
 DR SMART: SM00448; REC; 1.  
 DR PROSITE: PS0109; HIS\_KIN; 1.  
 DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.  
 DR Kinase; Complete proteome.  
 KW Kinase; Complete proteome.  
 SQ SEQUENCE 933 AA; 104605 MW; 5755C05F713E561D CRC64;

Query Match 25.3%; Score 171.5; DB 16; Length 933;  
 Best Local Similarity 33.9%; Pred. No. 1.4e-06;  
 Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;  
 QY 18 VLIVDDPDLNLIHKKIATGGISQTANNGEEAV-IHRDGGSSFDLILMDKEMPERDG 77  
 Db 811 ILVVDHPINRLLADQGLSGYCKTANDGVDALNVL--SKNHIDIVLSDVNNPNDMG 868  
 QY 78 STTKKLREMEVKSMTGVTSIADNEERRAFMEAGLNHCLAKPLTKDKI 126  
 Db 869 RLTIQRIQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDDVI 916

RESULT 10  
 Q8ZGR4 PRELIMINARY; PRT; 957 AA.  
 AC Q8ZGR4;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Two component sensor kinase/response regulator protein RcsC  
 DE (EC 2.7.3.-) (Sensor for ctr capsule biosynthesis).  
 GN RSCC OR YP01217 OR Y2971.  
 OS *Yersinia pestis*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=CO-92 / Biovar Orientalis;  
 RA MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Kariyshev A.V.,  
 RA Leatherell T., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,  
 RA "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM5 / Biovar Mediaevalis;  
 RX MEDLINE=22137863; PubMed=12142430;  
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
 RA Perry R.D.;  
 RL "Genome sequence of Yersinia pestis KIM."  
 RL J. Bacteriol. 184:4601-4611(2002).  
 DR EMBL; AJ414147; CAC90055.1;  
 DR EMBL; AE013898; AAM86522.1;  
 DR HSSP; P06143; I042;  
 DR InterPro; IPR003594; AtPbind\_Atpase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003661; His\_Kinase.  
 DR InterPro; IPR005467; His\_Kinase.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF00518; HATPase\_C; 1.  
 DR Pfam; PF00512; HiskA; 1.  
 DR PRINTS; PR00344; BCTRSENSOR.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00388; HiskA; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00091; PAS; 1.  
 DR SMART; SM00448; REC; 1.  
 DR TIGRfams; TIGR00229; sensory\_box; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 DR PROSITE; PS50113; PAC; 1.  
 DR PROSITE; PS50112; PAS; 1.  
 DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
 DR Kinase; Transferase; Complete proteome.  
 KW Kinase; Phosphorylation; Sensory transduction; Transferase.  
 SQ SEQUENCE 957 AA; 108591 MW; 0447A11F59100011 CRC64;  
  
 Query Match 25.3%; Score 171.5; DB 16; Length 957;  
 Best Local Similarity 35.1%; Pred. No. 1.5e-06;  
 Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;  
  
 QY 16 LNVLTVDLPLNLIHEKIIKAGGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERD 75  
 DB 835 LQILVVDHPINRRLADQLTTCYRVITANDGLDALVALNT--NTVDVLTVDVNPMD 892  
 QY 76 GVSTTKKLREMEVKSIMVGVTSLSADNEERRAFMEAGLNHCLAKPLTKDKI 126  
 DB 893 GYRLTERLURLNHNHFPPIIGVTANALAEKQRC-IEAGMDNCLSKPVTLDL 942  
  
 RESULT 11  
 Q9ANYO PRELIMINARY; PRT; 927 AA.  
 ID Q9ANYO;  
 AC Q9ANYO;  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DE Two-component system sensor protein.  
 DE XAC0494.  
 OS Xanthomonas axonopodis (pv. citri).  
 DE RSCS.  
 DE RSCS.  
 OS Vibrrio fischeri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=668;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=21142508; PubMed=11208780;  
 RA Visick K.L., Skoufos L.M.;  
 RT "A two-component sensor required for normal symbiotic colonization of  
 RT Euprymna scolopes by Vibrio fischeri."  
 RL J. Bacteriol. 183:835-842(2001).  
 CC -I- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE  
 CC KINASES.  
 CC EMBL; AF319618; AAG60694.1;  
 DR InterPro; IPR003594; AtPbind\_Atpase.  
 DR InterPro; IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro; IPR003661; His\_Kinase.  
 DR InterPro; IPR005467; His\_Kinase.  
 DR InterPro; IPR002570; Hpt.  
 DR InterPro; IPR001610; PAC.  
 DR InterPro; IPR000700; PAS-ASSOC\_C.  
 DR InterPro; IPR000014; PAS\_domain.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF02518; HATPase\_C; 1.  
 DR Pfam; PF00512; HiskA; 1.  
 DR Pfam; PF01627; Hpt; 1.  
 DR Pfam; PF00785; PAC; 1.  
 DR Pfam; PF00959; PAS; 1.  
 DR Pfam; PF00072; Response\_reg; 1.  
 DR PRINTS; PR00344; BCTRSENSOR.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00388; HiskA; 1.  
 DR SMART; SM00073; HPT; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00091; PAS; 1.  
 DR SMART; SM00448; REC; 1.  
 DR TIGRfams; TIGR00229; sensory\_box; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 DR PROSITE; PS50113; PAC; 1.  
 DR PROSITE; PS50112; PAS; 1.  
 DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
 DR Kinase; Phosphorylation; Sensory transduction; Transferase.  
 KW Kinase; Phosphorylation; Sensory transduction; Transferase.  
 SQ SEQUENCE 927 AA; 105590 MW; 9EB0659CB2E89F8B CRC64;  
  
 Query Match 25.2%; Score 171; DB 2; Length 927;  
 Best Local Similarity 33.9%; Pred. No. 1.6e-06;  
 Matches 43; Conservative 28; Mismatches 42; Indels 14; Gaps 4;  
  
 QY 16 LNVLTVDLPLNLIHEKIIKAGGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERD 75  
 DB 659 LSILIAEDNKINALVAQMFQRLGHKAIATENCK--VAIDKLKETHFDLIINDNHPVMD 716  
 QY 76 GVSTTKKLRE-MEVKSIMVGVTSLSA-----DNEERRAFMEAGLNHCLAKPLTKDKI 129  
 DB 717 GILATKIIRKIKISTVIFAYTANAFQKAHDN-----FLKAGANYVLTKPLOENDFIGA 770  
  
 QY 130 INQLMDA 136  
 DB 771 IKQYODA 777  
  
 RESULT 12  
 Q8PQ37 PRELIMINARY; PRT; 769 AA.  
 ID Q8PQ37;  
 AC Q8PQ37;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DE Two-component system sensor protein.  
 DE XAC0494.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RC MEDLINE=22022145; PubMed=12024217;  
 RX

```

RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergio F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL: AB011676; AM35383.1; -
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR002570; Hpt.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 2.
DR PRINTS: PD00344; BCTRSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 2.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 765 AA; 83699 MW; 87995CA40A787593 CRC64;

Query Match 25.1%; Score 170; DB 16; Length 769;
Best Local Similarity 30.1%; Pred. No. 1.5e-06;
Matches 36; Conservative 38; Mismatches 36; Indels 8; Gaps 3;

OY 14 KKLNVLDVDDPNLIIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLLMDKEMPE 73
DB 495 REVLIVDDPNLIIHVAQLAVLGFEDATDGEAA--LAMESTRYDMVFMQCPV 552

OY 74 RDGVSTTKKLEMEVKS-----MIVGVTSLADNEBERAFMEAGLNHCLAKPLTKDKI 126
DB 553 LDGYATRRWRANMETESGGRPVPIVAMTANAMAGDRERC--LAAGMDYLSKPVAREQL 609

RESULT 13
QBPFJN8 PRELIMINARY; PRT; 1364 AA.
AC QBPFJN8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component system sensor protein.
GN XAC2492.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergio F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL: AB011676; AM35383.1; -
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR002570; Hpt.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PD00344; BCTRSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 765 AA; 83699 MW; 87995CA40A787593 CRC64;

```

```

RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL: AB011887; AM37343.1; -
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR006189; CHASE.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR002570; Hpt.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF03924; CHASE; 1.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hiska; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 2.
DR PRINTS: PD00344; BCTRSENSOR.
DR ProDom: PD000039; Response_reg; 2.
DR TIGRFAMs: TIGR00229; sensory_box; 2.
DR PROSITE: PS50839; CHASE; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50113; PAC; 2.
DR PROSITE: PS50112; PAS; 2.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 2.
KW Complete proteome.
SQ SEQUENCE 1364 AA; 150282 MW; F7CCA3856B7E7165 CRC64;

Query Match 24.7%; Score 167.5; DB 16; Length 1364;
Best Local Similarity 29.2%; Pred. No. 5e-06;
Matches 35; Conservative 37; Mismatches 45; Indels 3; Gaps 2;

OY 17 NLLIVDDPNLIIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLLMDKEMPERDG 76
DB 1033 HLLIVDDSDNCVAGQILLEGAMVTVAHDGEQAVSTLKRAPNLFHLVLDVQMPVVDG 1092

OY 77 VSTTKKLEMEVKS--IVGVTSLADNEBERAFMEAGLNHCLAKPLTKDKIPLINQIM 134
DB 1093 YEATRLRQIPALASLPVIALTAGAFRPOEKA--LEAGNGFIAPFNVEELVTAIRHFL 1151

RESULT 14
QB5663 PRELIMINARY; PRT; 507 AA.
AC QB5663;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RCSC (Fragment).
GN RCSC.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB2000;
RX MEDLINE=95047557; PubMed=9829920;
RA Belas R., Schneider R., Welch M.;
RA "Characterization of Proteus mirabilis precocious swarming mutants;
RT identification of rsba, encoding a regulator of swarming behavior.";

```

```

RL J. Bacteriol. 180:6126-6139(1998).
CC -J- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AF071215; AAC82662.1; -.
DR HSSP: P06657; 2CHF.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; HSKA; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 2.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 2.
DR TIGRFAMs: TIGR00229; sensory_box; 2.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS00839; CHASE; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0113; PAC; 2.
DR PROSITE: PS0112; PAC; 2.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 2.
DR Complete proteome.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
FT NON_TER 1
SQ SEQUENCE 507 AA; 57326 MW; F5DA2EAD9C35DEF5 CRC64;

Query Match 24.6%; Score 166.5; DB 2; Length 507;
Best Local Similarity 36.4%; Pred. No. 1.9e-06;
Matches 39; Conservative 26; Mismatches 39; Indels 3; Gaps 2;

QY 16 LNVLIYDDPLNLIIHEKIIKAIGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERD 75
DB 397 LTVLIYDDHPINLLITDQKIKIGNTAEDGCCDALAFMQE--NHYDIITDVNPNNN 454
QY 76 GVSTTKKLEMEVKSMIVGVTSLADNEEERRAFMEAGLNHCLAKPLT 122
DB 455 GYLQATTVRELSTPIIGVTANATAEEKQRC-IDAGMNDVCVKDVS 500

RESULT 15
ID Q8P883 PRELIMINARY; PRT; 1364 AA.
AC Q8P883;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component system sensor protein.
DE XCC2360.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 529;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarette G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira L.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities."
RL Nature 417:459-463(2002).
DB EMBL: AE012344; AAA41638.1; -.

```

```

DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR006189; CHASE.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF03524; CHASE; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; HSKA; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00785; PAC; 2.
DR Pfam: PF00989; PAS; 2.
DR Pfam: PF00072; response_reg; 2.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 2.
DR TIGRFAMs: TIGR00229; sensory_box; 2.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS00839; CHASE; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0113; PAC; 2.
DR PROSITE: PS0112; PAC; 2.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 2.
DR Complete proteome.
KW SEQUENCE 1364 AA; 150167 MW; 48F1C5FED9710316 CRC64;

Query Match 24.1%; Score 163.5; DB 16; Length 1364;
Best Local Similarity 30.3%; Pred. No. 1.1e-05;
Matches 36; Conservative 36; Mismatches 44; Indels 3; Gaps 3;

QY 18 LVIVDDPLNLIIHEKIIKAIGISQTANNGEEAVIIHRDGGSSFDLILMDKEMPERDGV 77
DB 1034 LLLVDDSEINCEVAQRILEGEGAMVTVAHDGEQAVNTLKRAPDLFHLVIMDVQHPVVDGY 1093
QY 78 STTKKLEME-EVKSM-IVGVTSLADNEEERRAFMEAGLNHCLAKPLTKDKIIPLINQM 134
DB 1094 EATRLRQIPSLASLPVIALTAGAFRPQOEKA-LEAGMNGFIAPKFNVEELVTAIRHFL 1151

```

Search completed: August 13, 2003, 19:12:15  
Job time : 39 secs

US 099043890KP1



Creation date: 03-29-2004  
Indexing Officer: SMOHAMMED - SUAD MOHAMMED  
Team: 1600PrintWorkingFolder  
Dossier: 09904389

Legal Date: 10-25-2003

No.	Docode	Number of pages
1	SRNT	94

Total number of pages: 94

Remarks:

Order of re-scan issued on .....